

Inquest or enquiry into research?

The near-collapse of the British research enterprise has become a public issue. The extra funds announced this week may be a palliative, but cannot be a cure. The House of Lords has the luck to look closer.

THE endless preoccupation of the British academic community with the financing of basic research is in some respects a parochial matter. There is some force in the jibe, heard elsewhere, that British science would be in better shape if people spent less time worrying about the machinery of research support and more time at the bench. There are two sets of reasons why more people do not follow this simple course. First, for quite extraneous reasons, structural changes are being forced on British establishments, research institutes and universities alike, by the need to match spending with the nominally fixed but depreciating income which the government provides. Second, the British scientific community has been brought up against the hard truth that the quality of British research has sharply declined in the past few years in ways most simply but not fully explained by the enforced transformation of the research environment. People elsewhere may regard the present upheaval in British research, where the accountants are scouring the field for projects that can be brought to an end, as a salutary illustration of what is bound to happen when economic failure makes it impossible to sustain a research enterprise on the scale that tradition would dictate. To the British, no doubt, the present crisis seems more like a battle for survival.

The result is that the question of the sufficiency of research support has become, no doubt fleetingly, a matter of public interest. Earlier this week, there was an hour-long television programme built around a demonstration that British contributions to the international literature, especially in fields such as physics, declined rapidly in the late 1970s. Next Wednesday (21 November) the House of Lords, still the most urbane of democratic institutions, plans a debate on the subject. The issue then, as it has been for the past several years, will be that of reconciling the British government's view that it can afford no more for research ("£500 million is a lot of money" in the Prime Minister's memorable phrase) with the research community's frustrated conviction that it is hamstrung by the general lack of resources. Both the Prime Minister and the community are right. Will the House of Lords have the wit to find common ground between apparently conflicting perceptions?

Corrosion

The impoverishment of the British research enterprise is not in question. The extra funds for the research councils and for university equipment made available earlier this week (see p.185) may take the edge off some anxieties, but they will not touch the essence of the problem — the general decline of enthusiasm for research and the erosion of young people's belief that energy and flair can still bring excitement. The rot set in more than a decade ago and has been corroding the character of the research enterprise ever since. The analysis of contributions to the international literature by people of different nationality elaborately displayed last week by John Irvine and Ben Martin of the University of Sussex (*Foresight in science: Picking the winners*, Frances Pinter, London) catalogues the symptoms of decline. Between 1973 and 1980, the British share of physics publications declined from 7.7 per cent to 5.9 per cent of the total, for example. The causes are necessarily more elusive. The system has become too rigid, with the result that even talented people with a good idea of what should be done, and who may be able to support themselves in the research enterprise, are not now able to recruit help on a scale that

will allow them to move quickly. In a competitive world, the best are destined to be second. They are also certain to be distracted by the extraneous demands their institutions make of them — to think of spinning a penny from industrial collaboration, to sit on a committee adjudicating the virtues of somebody else's research, to draft a prospectus that will offer academic or research services to unfamiliar customers. Sir Keith Joseph was right, earlier this week, to hope that his modest extra few per cent will be spent selectively. British academic research needs a few striking examples of projects done well to lift its spirits against better times. But money by itself will not put things right.

Impoverishment

The impoverishment of the public purse is also plain for all to see. The British treasury minister, Mr Nigel Lawson, left those who heard him earlier this week in no doubt of that. The British government has won its battle against the inflation of the 1970s, but is clearly puzzled that the result has not been the return of competitiveness in commercial markets or the resurgence of investment in new industrial enterprises. The best it can hope for now is that the trends will right themselves if it sticks to its present policies. It is worth noting that those who disagree with these economic policies do not advocate that there should be more spending on research. They would build roads, airports and modern hospitals instead. The economic box in which Britain is constrained does not offer a plausible escape along a path strewn with larger budgets for research. Even the Prime Minister's political opponents would probably, in present circumstances, be saying that £500 million is a lot of money.

So the remedy for the research enterprise must lie in the way in which the system of support is organized. This has been clear for the past three years, at least since it first became apparent that the Agricultural and Food Research Council was so heavily committed to institutes for the conduct of research that had seemed desirable more than a decade earlier that it had too little left for the pursuit of more immediate goals. Since then it has become apparent that the same fault runs through the whole system, the research councils and the universities together. And, one by one, these institutions are coming to appreciate that they will have to change in such a way that their resources are concentrated only on the most excellent projects. The time for the second-rate has gone, yet the second-rate survives by inertia, the persistence of too much politeness in the judgements people make of others' projects and the general sense that venerable institutions deserve preservation for their own sake. The academic enterprise itself must take some of the blame for failing to recognize the need for change (now underestimated). The government's fault (Sir Keith Joseph's £24 million notwithstanding) is its assumption that a national resource as important as the complex of academic institutions and the civil research establishments will of its own accord find the best strategy for survival if exposed to enough pressure for long enough. What if, by mismanagement or misjudgement, it elects to collapse under its own weight instead?

Those who take part in next week's debate in the House of Lords therefore have a delicate balancing act to perform if they wish to be convincing. They have to accept the government's case, that there is hardly any extra money, but also the claims of the



research enterprise that the pressures to which it is exposed are already dangerous. The trick must be to encourage and even facilitate change in circumstances that are conducive to the opposite. The method must be to provide a host of different illustrations that change is not merely possible but likely to be beneficial. Taking an axe to some parts of bureaucracy would be a start. Too many people at present spend too much time on deliberations likely in the end to be pointless because there are no funds with which to back the outcome. The Natural Environment Research Council, its chairman's recent defence apart, seems a natural place to start. Simplifying the problem of managing expensive fixed facilities (telescopes, accelerators and the like) by wrapping them in a large bundle and putting them at the disposal of European scientists is another. Sweeping away the restrictive practices which at present ensure that large sums of money dedicated to graduate training are parcelled out to research supervisors, not to the young who will do the research, is a mistake that needs rectifying anyway, but whose remedy would change the rules of what has become a depressing game. There is, in other words, no shortage of things that might be done. But nothing will count for as much as the wit of the grant-making agency that is able to find an imaginative project, one already canvassed widely, and to back it to the hilt. The selectivity that everybody (the minister included) now asks for consists of being positive, not of finding new ways of saying no.

The House of Lords, being what it is, is also certain to take up the question of central administration. Four years ago, its own select committee tried to revive the appointment of a minister (part-time) with responsibility for science and technology, and was given a dusty answer by the present government. But the principles then apparent apply even in present circumstances. Although much of what is wrong with the academic research enterprise stems from faults within the system, it is also plain that the general coordination of publicly-supported research is inefficient. The Rothschild doctrine that government departments should become customers for research services through the agency of energetic chief scientists has long since collapsed under the reluctance of departments to spend money on any but problems that need solving by tomorrow. And defence research remains outside the system, and a law unto itself. The consequence is that the academic research enterprise is too isolated from the rest of what passes for research and, perhaps more important, that the British economy benefits much less than it should from what the government spends on all kinds of science. The government may not give the House of Lords its minister, but can it reasonably deny that there should be a mechanism for making better use of what is, after all, a lot of money? □

What price drugs?

New categories of drugs for which British patients must pay set dangerous precedents.

WHEREVER there is health insurance, there are problems about the prices of modern pharmaceuticals. In most advanced societies, the cost of drugs exceeds 15 per cent of the cost of health services and, until recently, has also been a rising proportion in most places. That is why, throughout Europe, there have grown up elaborate systems by means of which governments attempt to control the cost of medication. The latest development is in Britain, where last week the minister responsible for the National Health Service, Mr Norman Fowler, announced a novel scheme for excluding from the services provided free by the health service two categories of medicines that are widely used — certain non-prescription medicines including aspirin but also some near-cosmetics which, it is suggested, people might in future prescribe for themselves and, second, tranquillizers and sleeping pills based on the diazepam structure which, the ministry has decided, are too often used as needless palliatives. Physicians will not of course be forbidden to prescribe these drugs, but patients will have to pay the full cost, not merely the standard fee for the servicing of a prescription by a pharmacist. There has been some grumbling on

both counts, although that attaching to the first category of non-prescription drugs is insubstantial. The second class of drugs raises more interesting questions.

On the face of things, the obvious objection to discrimination between different categories of medicines is that it is an interference with medical practice. Schemes such as these, however, are used to control the cost of pharmaceuticals in both Sweden and France. The general principle seems to be that the drugs of chronic sickness or those required in serious illness or which are expensive should cost the patient nothing or very little, but the non-prescription medicines should be paid for as if, indeed, they were cosmetics. Elsewhere, in Switzerland as well as France, there are also approved lists of drugs that may be supplied within the terms of public health insurance, with a general understanding that approval will not be given to drugs that are notoriously less economical in use than others also on the market. Unsurprisingly, Switzerland is more permissive than France.

Nobody can reasonably complain at the British government's intention to exclude non-prescription drugs from the lists of those supplied free, although the saving may not be as great as expected when patients learn to persuade their physicians to prescribe more powerful (and more expensive) substitutes. The intended exclusion of the diazepam is more tricky because the category is so specific. The fact that the principal supplier of the drugs sold in Britain as Valium and Mogadon is the Swiss-based company Ciba-Geigy gives the decision a sense of political realism but may yet be a dangerous precedent. And while there seems ample evidence that the use of prescription tranquillizers is much more common than it should be, with some people as dependent on their pills as on their circadian clocks, what amounts to a decision that these relatively harmless materials should not be used for the treatment of recognizable conditions, while not a deliberate interference with the freedom of physicians, is a needless constraint on their way of working. A more generally drawn category of drugs qualifying for partial reimbursement would have been a wiser choice.

But why charge for drugs at all? That is the common complaint of the pharmaceutical manufacturers, who argue that attempts to regulate the cost of medicines can only reduce the pace of innovation and thus militate against the welfare of patients. The argument has some force, and is also accepted by most governments, which are unwilling to hazard the prosperity of their manufacturing industries. But most European governments also operate schemes for the regulation of the prices of drugs supplied within the public health services. In Britain, for example, the principle is that profits should be regulated in relation to the capital employed in research, development, manufacturing and marketing. What the pharmaceutical manufacturers do not acknowledge openly enough is that these arrangements are open to abuse. A manufacturer with a strongly-selling drug can set off against the income expenditure of the wildest kind — not merely the now well-publicized extravagances by which physicians are persuaded to prescribe particular drugs but also ill-conceived programmes of research and development. If the successful drug is one in great demand, and if there are no obvious competitors, the manufacturer has what was once described, in another context, as a licence to print money. Governments cannot equitably let such a state of affairs persist. The pharmaceutical industry, which is prone to howl whenever governments interfere, could usefully spend more of its energy trying to devise schemes for regulation which are equitable to both sides.

The issue is one on which the European Commission could usefully take a stand. Most community countries have public health insurance, either a national health service (sometimes called socialized medicine) as in Britain or compulsory private health insurance (which comes to the same thing) as in West Germany. Paying for non-prescription drugs is sensible, making chronic drugs free of charge is humane. Price regulation should not bear on specific drugs or on the industry as a whole, but on classes of drugs for the treatment of particular conditions which can be expected to proliferate with the passage of time. And prices should be uniform. □

UK research budget

Government promises redirection of extra funds

THE British government has found an extra £24 million for the support of academic research in the next financial year and that succeeding. This was announced on Monday this week by Sir Keith Joseph, Secretary of State for Education and Science, who said that the science budget (used chiefly for the support of the research councils) would be increased by £14 million and that the University Grants Committee would have an extra £10 million to spend on equipment (compared with the present £90 million).

Sir Keith said on Monday that his objective has been, within a budget for education and science which is essentially unchanged, to redirect funds towards research "which is in urgent need". One measure of his commitment is that the extra funds have been found by increasing the contributions well-to-do parents will be expected to make towards the cost of their children's higher education. The minimum maintenance grant for students from better-off families is to be abolished from the next academic year, and parents will be expected to contribute towards tuition fees up to a maximum of £525 a year.

Within the science establishment, the chief beneficiaries are likely to be the science and medical research councils, which had been expecting to contribute towards the cost of reorganization elsewhere in the coming financial year. But Sir Keith said on Monday that it would be for the Advisory Board for the Research Councils to recommend the precise division of the spoils. The extra £10 million for university equipment is to be spent "in a few carefully-selected centres of research".

Sir Keith's conviction of the urgency of the need for extra research funds stems, by his statement on Monday, from "talking to people" but also from reading the report of a committee under Sir Jack Lewis, the Cambridge physical chemist, that detailed the proportions of high-quality research applications being turned down for lack of funds by the Science and Engineering Research Council (*Nature* 310, 267; 1984).

Whether the extra funds will satisfy the research councils or their needs is another matter. The Science and Engineering Research Council has been saying that its budget would have to increase by £70 million by the end of the decade simply to sustain present activities. It also became plain this week that there is no general understanding that the research councils should be protected against the effect of currency fluctuations.

The universities are likely to be equally disappointed by the fixed recurrent budget

which, after allowing for the upward drift of salaries means a reduction of resources by between 0.5 and 1.0 per cent a year.

SERC annual report

Council puts on brave face

THE UK Science and Engineering Research Council (SERC) gave no sign last week, on the publication of its report for 1983-84, of the turmoil behind the scenes, as a network of committees embarks on cutting the council's activities to fit within its budget. Professor John Kingman, the unruffled chairman, highlighted British involvement in the Infra-Red Astronomy Satellite, the discovery of intermediate bosons at CERN and SERC's teaching company scheme.

Since its inception in 1976, the scheme, involving a collaboration between universities and industry whose results tend to be directly measurable by increased manufacturing efficiency, has grown considerably. By 1988, the number of individual schemes will have increased to 235 at an annual cost of £11.2 million, of which £3.7 million will be borne by the companies involved.

The teaching company scheme is one of the successes of the council's engineering board, whose share of council spending is set to increase from 18.1 per cent in the financial year 1979-80 to 28.1 per cent in

What to abandon?

How likely is it that Britain will pull out of the European Organization for Nuclear Research (CERN) — so saving SERC its £32-million annual subscription? One scenario being canvassed last week was that the SERC committee reviewing CERN membership, headed by molecular biologist Sir John Kendrew, would recommend withdrawal but that the government would reject the proposal because of the political symbolism of CERN as one of the few working European institutions.

But this is not very likely, says John Kingman, SERC chairman. He offered another more frightening scenario last week. Kendrew might recommend, hypothesized Kingman, SERC should stay in CERN. But the SERC council might have to overturn the Kendrew recommendation and get out of CERN anyhow. Kingman thinks the political objections to leaving CERN would be balanced by the advantages of using some of the money released to join a new European project — the European Synchrotron Radiation Source.

Robert Walgate

In late September, the grants committee offered its opinion that a continuation of this pattern would soon make it necessary to think of closing "one or more" universities. Sir Keith said on Monday that he did not accept this outcome as inevitable, and that the study of academic efficiency mounted by the Committee of Vice-Chancellors and Principals could yet show that there were substantial economies to be made in higher education. □

1987-88. Over the same period, the percentage expenditure of the other boards will have uniformly decreased. Concern at these trends may have been in the mind of the chairman of the nuclear physics board, Professor D. Colley, when he remarked last week that total support for applied research from a variety of sources far outweighs that for basic science, provided solely by SERC. Thus further shifts towards applied research by SERC would be out of proportion to the damage they would cause to basic research. Professor Kingman had earlier emphasized, however, that engineering research was not synonymous with applied research.

Another shift in SERC's priorities is marked by the approaching end of an era in which large facilities have been constructed. All future large facilities are expected to be developed on an international basis.

As a result, it is expected that funds may be released towards recurrent expenditure on project grants. Such funds are likely to be badly needed. Professor Kingman repeated a remark made by the chairman of the Advisory Board for the Research Councils, Sir David Phillips, to the effect that the level of UK scientific activity seems likely to decrease by 25 per cent over the next decade.

Professor Kingman emphasized last week that, although final responsibility rests with the council, use would be made of peer review. The process would be "loosely coupled" with that of the University Grants Committee, which is also embarking on a re-examination of its support for science. Professor J. Cadogan, chairman of the science board, emphasized that, although four months seems a short time to reach such decisions, the council has been reviewing its programmes for a variety of reasons over recent years, and so is reasonably equipped for such a task.

When asked what amount of money would solve his problems, Professor Kingman said that SERC's proportion of the gross national product peaked in 1971-72 and has declined ever since, implying a current lack of £25 million or so. But just a guarantee from the government of continued support on the present scale, he said, would leave SERC in a much stronger position.

Philip Campbell

US election

Business as usual for science

Washington

If the results of last week's US presidential election contained few surprises, they also imply that the administration of science will continue unchanged. This has been a good year for incumbents in the Congress. There will be fewer than 50 new faces in the 435-member House of Representatives and only six in the Senate, which means little change in the committee power structures of Congress.

The most noticeable shift will be the move of Representative Albert Gore (Democrat, Tennessee) to the Senate. Gore, who chaired the investigations subcommittee of the House Science and Technology Committee, took an unusual personal interest in scientific issues, initiating well-publicized investigations of biotechnology, human genetic engineering and university-industry relations (including the multi-million dollar deals between Washington University and Monsanto and between Massachusetts General Hospital and Hoechst), biomedical research fraud, organ transplants and hazardous waste.

His staff may be forgiven the minor exaggeration that Gore's move to the Senate marks the "end of an era" for the Science and Technology Committee; it is indeed hard to see who on the committee is equipped, by temperament or knowledge, to follow in his footsteps. Curiously, the two other members of the House who also ran successfully for the Senate — Tom Harkin (Democrat, Iowa) and Paul Simon (Democrat, Illinois) — were also on the Science and Technology Committee, although neither held a chairmanship.

Representative George Brown (Democrat, California), a leading voice for science in the House, was re-elected by a surprisingly wide margin, given his recent troubles in his increasingly Republican district. Brown has led the fight for competitive grants for agricultural research and has been a key opponent of the testing and deployment of antisatellite weapons in particular and Star Wars in general.

Brown has not yet decided whether to seek a subcommittee chairmanship on the Science and Technology Committee or retain his position as chairman of the research subcommittee on the Agriculture Committee. He has moved back and forth before; with the farm bill up for renewal this year and the prospect of a major push to expand the competitive grants programme, however, he may find a compelling reason to stay put.

In the Republican-controlled Senate, the major change in prospect is on the Foreign Relations Committee, which would have the first crack at any arms-control treaties submitted for Senate ratification.

The chairman had been moderate Republican Charles Percy of Illinois, the

man who, incidentally, fought the valiant but unsuccessful fight to overrule an Energy Department scientific advisory panel and award the contract for a new electron accelerator to Argonne National Laboratory, which happens to be in Illinois. He was defeated in part through the good offices of the National Conservative Political Action Committee, which decided it would prefer an extra Democrat in the Senate than Percy in command of foreign policy.

Next in line for the committee chairmanship is staunch conservative Jesse Helms (Republican, North Carolina), who has expressed strong reservations about arms control and who holds the distinction of being the only senator to support the

Argentinian military regime in the Falklands War. Helms promised during the campaign to stay put as chairman of the Agriculture Committee — an important point in his home state, always worried about losing its tobacco subsidies. But Helms is said to be under strong pressure from the right-wing of the Republican party to take the foreign relations post.

There have been no strong signals about the course the Reagan administration will take towards science and technology in its second term. It is safe to assume that the National Science Foundation and the National Institutes of Health will be protected as they have been; that defence research and development will continue to grow faster than civilian; and that, if Reagan hews to his promise to reduce the deficit without raising taxes, that hopes for big science in the form of accelerators or planetary missions will have to be tempered.

Stephen Budiansky

Thermonuclear fusion

Japan's tokamak and after

Tokyo

THE assembly of Japan's giant tokamak, JT60, was completed this week and final sub-system tests are under way. If all goes well, the whole system will be in operation in the spring, plasma heating experiments will start the following year and break-even will be achieved in 1987.

The JT60 project, run by the Science and Technology Agency's Japan Atomic Energy Research Institute, is on much the same scale as the US Tokamak Fusion Test Reactor (TFTR) at Princeton University and the Joint European Torus (JET), but is running a little behind them: TFTR successfully produced its first test plasma almost two years ago and the assembly of JET was completed last year.

The project is not cheap: the budget for this year alone is around ¥29,000 million (£96 million) and total expenditure is expected to be considerably higher than that at Princeton. Most of the equipment is built by the giant engineering company Hitachi, which has one of its manufacturing facilities conveniently close by, at Hitachi city.

Despite the overall similarities between JT60, TFTR and JET — all are of tokamak configuration with an external radius close to 3 metres — JT60 is not intended to follow exactly the same research path as the other two designs. No attempt is to be made to create a deuterium-tritium plasma, the fuel that a commercial reactor would have to use. The energetic neutrons produced in such reactions would make the tokamak structure radioactive, requiring all maintenance reactions to be carried out remotely. Instead, only hydrogen and deuterium fuels will be used, and the aim will be to study the confinement of plasmas that would produce break-even conditions if tritium were present. At both TFTR and JET, there will be a switch to tritium and

deuterium fuels, and remote handling, towards the end of the decade.

Plasma heating in JT60 is to be achieved by a combination of 14 neutral beam injectors, firing high-energy particles into the plasma, with microwave heating in the 2 GHz range. The effect should allow much longer heating pulses to be produced than in TFTR, which relies more on ohmic heating. JT60 will also incorporate a poloidal divertor to skim impurities from the edge of the plasma, a sophisticated piece of technology that had not been developed when TFTR was being planned.

After JT60 will, in theory, come the experimental test reactor, which will aim in the 1990s for tritium-deuterium self-ignition, a burn of more than a hundred seconds and the self-sustaining breeding of tritium in the lithium blanket surrounding the reactor. But like the United States and the European Community, Japan is worrying about just how much longer it can maintain the enormous expense of an independent fusion reaction programme that inevitably duplicates some of the effort made elsewhere.

Japan already has a long-term cooperation agreement with the United States in fusion research and there is every likelihood that some joint technical workshops — and perhaps exchange of engineers — will take place as both Japan and the United States move to the design stage of the next generation of fusion machines. But whether closer cooperation between the United States, Japan and Europe is possible on a large scale is more doubtful: discussion of the merits of a joint fusion materials test facility has been dragging on for more than a year (see *Nature* 306, 215; 1983) without the European Community being able to commit itself, even though both the United States and Japan have expressed enthusiasm.

Alun Anderson

European Community

Programmes hanging fire

Brussels

THE fate of eight European Community long-term research and development programmes still remains unresolved after the Research Council meeting on 6 November.

This was the third attempt to take a decision on certain projects, such as the radiation protection training programme, biotechnology and non-nuclear energy, whose second research and development programme expired almost eighteen months ago.

Part of the problem has been the lack of a decision on the UK rebate, now provisionally resolved. Also, a large share of the 1985 research budget has already been approved for the Esprit (144.5 million European Currency Units (ECU); 1 ECU = £0.60) and Joint Research Centre (215.5 million ECU) programmes, which are favoured by the larger countries of the Community, the United Kingdom, West Germany and France, which have also proposed cuts in funds for the smaller sectoral programmes. This has left about 150 million ECU for 1985 to be shared out among the other programmes.

These constraints mean that some of the programmes have to be whittled down or phased in gradually. The question is which ones? Little progress appears to have been made on the positions held by the ministers as they prepared to discuss the issue at their meeting. The following are the programmes involved.

- Multiannual training and research in the field of radiation protection. The European Commission's original proposal of 93.4 million ECU for 1985-89 could be reduced by a third.

- Fundamental technological research aimed at promoting industrial competitiveness in industries other than high technology (known as the BRITE programme). The Commission proposed 170 million ECU for this programme, but the United Kingdom and West Germany want a smaller sum.

- Stimulation of cooperation and exchanges in the field of scientific and technological research, for which the commission has proposed 40 million ECU for 1985-88.

- Biotechnology, for which the commission originally proposed 88.52 million ECU. This programme originally had the backing of the French but more recently both France and the United Kingdom have proposed a reduction.

- Non-nuclear energy involving the further development of solar energy, biomass, wind energy, geothermal energy, energy saving and the use of solid fuels and possibly research in the field of hydrocarbons. Unlike the smaller countries, the United Kingdom, France and West Germany want a considerable reduction in the proposed commission budget of 379 million ECU for

1984-87.

- Reactor safety, consisting of a number of cost-shared projects to which the Community would, according to the Commission, contribute 24.9 million ECU over four years.

- Nuclear fusion, for which the Commission had proposed 790 million ECU over five years. This includes funds for the Joint European Torus at Culham in the United Kingdom and the setting up of a tritium-handling laboratory.

- Radioactive waste management programme to continue work under the current programme which expires at the end of the year. Ninety-two million ECU have been proposed by the Commission for this programme but, as in the case of radiation protection and reactor safety, no funds have been earmarked in the draft 1985 budget pending the adoption of detailed proposals.

Although no decisions were made on 6 November, research and development commissioner Vicomte Etienne Davignon managed to extract a commitment from ministers to come to an agreement by the next scheduled research council on 19 December. The decisions taken by the budget council on 29 November may make matters clearer by then.

The council has also become involved in the debate over the controversial siting of the European Synchrotron Radiation Source (ESRS) at Grenoble which Germany and France had agreed together (see *Nature* 25 October, p.695 and 1 November, p.3).

Although the Community is not directly concerned, Denmark insisted on a discussion at Community level, at which point Italian research minister Luigi Granelli declared that Trieste was also interested in ESRS, and that ties with Yugoslavia would provide a useful link with an "outsider" country.

The issue will appear on the agenda of the next research council, by which time the report of the standing committee should be ready. In the meantime the European Commission will also consider the possible involvement of the Community in ESRS.

Vicomte Davignon also presented research ministers with a project for inter-governmental videocommunications, originally proposed by the French in February. The link-up on a bilateral basis of Frankfurt, London, the Hague and Rome with Paris was to be tested this week. The first demonstration will be followed by a second link-up between the heads of state attending the European Council in 1985 and their advisers and cabinet ministers back home. The Commission intends to make a proposal to the council on the implementation and funding of the videocommunications programme at the beginning of next year.

Anna Lubinska

Europe

Cooperation needed

Brussels

IN addition to trying to encourage "solidarity" among governments of the European Community, research and development commissioner Vicomte Etienne Davignon was last week exhorting industry and researchers in Europe to cooperate more in the research and development of new technologies rather than continue on the path of "technological nationalism".

He was addressing a conference organized by the Community's economic and social committee aimed at bringing together for the first time all the sectors affected by the introduction of three new technologies: automation, information technologies and biotechnology.

Predictions for information technology and biotechnology are ambitious. Biotechnology, according to US consultants, could have a world market worth \$27,000 million by 1990, while information technology could be worth \$500,000. Meanwhile, the Community's trade deficit in information technology almost doubled between 1979 and 1982, reaching \$15,000 million.

Europe has a world market share of 10 per cent in the information technology field and only a 40 per cent share of its own market. But intra-European, interdisciplinary and industry/research cooperation could change that. The conference, which included trade union representatives, voiced concern about the impact on jobs of the introduction of new technologies.

Because there would be job losses, more training and retraining would be necessary, although the present unemployment problem was not blamed on the introduction of new technologies to any great extent. While it was mostly believed that new technologies would only provide a small percentage of the new jobs required, Umberto Colombo, president of the Codest committee and chairman of the Ente Nazionale di Energie Alternative, said that a study by the European Nuclear Energy Agency had identified 200 new job categories that might generate 3 million jobs over ten years.

But research and development requires money, which is precisely the area in which the Community is having problems. As Gerd Muhr, new chairman of the economic and social committee, pointed out, one US company spends on research in two months what the Community spends in a year on research, information, innovation and industrial policy put together. Only 0.5 per cent of the total Community budget was set aside for these areas in 1984. Funds for research in 1985 risk further curtailment.

Anna Lubinska

UK agricultural research

Disquiet over exclusive deal

GRUMBLES persist about the agreement between the British Agricultural and Food Research Council (AFRC) and the commercial company which now enjoys first refusal of innovations arising in its research laboratories. The company, Agricultural Genetics Company Ltd (AGC), was formed in July 1983 after more than a year of negotiation, during which the terms of the agreement with the council played a crucial part.

There are three categories of complaints, most common of which is that the exclusivity of the arrangement may impede rather than assist the transfer of new agricultural technology to the private sector. More particularly, there are fears that the deal may dissuade other companies from collaborating with AFRC, while the arrangements for financing AGC are held by some to have been inequitable.

AGC was originally intended as an agricultural analogue of Celltech, formed in 1980 with partial government support (40 per cent of the £12 million launch capital was provided by the British Technology Group) and now active in medical biotechnology (chiefly monoclonal antibodies). Celltech at first enjoyed a right of first refusal on Medical Research Council innovations which was terminated after three years.

Like Celltech, AGC was founded with backing from BTG, which provided £200,000 of the launch capital of £750,000. The agreement between AGC and AFRC,

described in outline in a paper circulated last month to the council's research staff, says that the company will have the "first option" to develop and market new ideas arising at council establishments in the fields of non-conventional plant breeding, microbial inoculants and biological control.

The right will be restricted to research at the East Malling Research Station, the Glasshouse Crops Research Institute, the John Innes Institute, the Plant Breeding Institute, the Rothamsted Experimental Station and the Unit of Nitrogen Fixation at the University of Sussex. The council estimates that some 30 per cent of the work of the institutes is covered by the agreement with AGC. Conventional plant breeding, excluded from the agreement, is at present exploited through the National Seed Development Organisation.

One common grumble is that AGC will have up to six months to decide whether to exploit an innovation arising within one of the six establishments. The council's document says that the delay will "usually" be less but, even so, it has promised members of staff that publications delayed in this way will nevertheless count in deciding promotion claims if unpublished manuscripts are "accompanied by evidence such as patent applications".

A more serious cause of anxiety is that the institutes covered by the agreement may not embark on collaborative research with other companies without first consult-

ing AGC which, it is said "in most cases is likely to consent". Given the notorious reluctance of research directors to disclose their plans to others, however, demand from third parties may henceforth be small.

The agreement also provides for the company to place research contracts with AFRC establishments both for the further development of indigenous innovations and those suggested by the company. AFRC expects that the value of this contract work will amount to £2 million a year in the next few years.

This modest income is one of the attractions to the council, which has also been given an undertaking that it may keep royalty and other income arising from its projects. But there is no assurance that the Treasury will not cut back on AFRC's total budget to allow for income earned. The council appears also to have won agreement that its establishments will be able to hire short-term staff to help carry projects through.

The commercial arguments about the financing of the company are different, and centre on the belief of some venture capital companies that they should have been given a fuller opportunity to participate at an earlier stage, and that the initial subscription of capital (£750,000) is incommensurate with the potential value of the enterprise.

Since the formation of the company, the Rothschild biotechnology fund has subscribed a substantial amount towards the capital of an expanded company, while Morgan Grenfell, the London merchant bank, is this week arranging a private placing of shares in AGC Ltd that will increase the capital of the company to £15.2 million.

The three founding investors will end up owning two-thirds of the company (roughly 5 per cent of which will belong to the chief executive, Dr Roger Gilmour). New entrants will have paid roughly four times as much for each of their shares as will those entering on the ground floor. Rothschilds will have made a better bargain.

John Maddox

Technology by contract

THE agreement between AFRC and AGC is said to reflect the research council's belief that it is the most effective way of transferring new agricultural technology to British industry. The council also says (in its staff paper) that the arrangement is "designed to give AGC a competitive advantage".

The agreement foresees three types of collaboration:

- Promising lines of research will be offered to AGC, which will have six months to decide. AGC will hold property rights (in return for a royalty on eventual sales) and will meet the cost of further development, usually at council laboratories.
- AGC will support contract research at council laboratories on its own initiative or in response to laboratory suggestions, meeting the full cost and owning such rights as may accrue.
- Research council institutes are not precluded from collaboration with third parties provided AGC consents (which it must decide within 30 days). AGC will seek collaboration with third parties for mounting research contracts. AGC may

seek grants from non-commercial sources, including foundations, provided that commercial exploitation by AGC is not precluded.

The agreement also stipulates that if AGC adopts a council innovation but does not prosecute it energetically, the rights will revert to the council, which will also be represented on the board of the company. The agreement would come to an end if ownership of AGC were not at least 60 per cent British.

Rules on confidentiality are said to be those which already apply, although the council's document says that these will have to be "strictly adhered to". "Restrictions on communications" will not apply within the council's staff "or with academics in related fields provided that commercial value is not diminished".

Day-to-day management of the agreement will rest with institute directors, but the council's headquarters will be involved on legal matters such as the negotiation of royalties. The agreement is for five years in the first instance, but is expected to be renewed for at least two further periods of that length. □

Demonstrable success

Mr William Coates, a member of the technical staff of the Royal Institution, London, since 1948 and the man in charge of demonstrations at the institution's Friday evening lectures (called discourses), exchanged roles with his director, Sir George Porter, last Friday evening (9 November). Coates had turned 65 earlier in the week, but intends to keep working for at least the next few years.

Coates explained that when he began work at the institution, it was required that technical staff should wear brown laboratory coats, while qualified scientists were dignified in white. Last week, he and Porter wore dinner jackets. □

Biological weapons

US Army's plan blocked

Washington

A US Army plan to build a \$1.4 million facility for testing "aerosol toxins" has been blocked by a single senator, who noticed the project buried away in a normally routine request by the Army to reallocate a total of \$140 million already appropriated for other projects.

While ostensibly a part of the Army's efforts to develop defences against chemical and biological warfare agents, the aerosol laboratory would, according to Senator Jim Sasser (Democrat, Tennessee), provide a "potential capability to test offensive biological toxins". The testing or possession of biological warfare



agents for other than prophylactic, protective, or other peaceful purposes is prohibited by the 1972 convention outlawing biological warfare.

The United States maintains an active programme of research on biological defence. Most of the work takes place at the US Army Medical Research Institute of Infectious Diseases at Fort Detrick, Maryland, the former biological warfare establishment. Its budget of roughly \$40 million a year is devoted mainly to the development of vaccines and other medical defences. Dugway Proving Ground in Utah, where the aerosol laboratory was to have been built, deals with the development and testing of defensive hardware and protective gear. Fort Detrick is totally open; Dugway's work is classified.

The Army proposal came as part of a "reprogramming" request sent to the House and Senate Appropriations Committees. Under this procedure, funds appropriated for one specific purpose can be transferred to another without full congressional approval. In the Senate, the concurrence of the ranking majority and minority members of the relevant appropriations subcommittee is normally all that is required. Sasser had earlier approved a

\$66 million reprogramming request that contained the aerosol laboratory, but after learning of the project, he withdrew his approval and also blocked a pending \$7 million request for "toxic agent test support facilities" at Dugway.

According to the reprogramming request submitted by the Army, the aerosol laboratory would be used to test biological "protective gear and detection/warning equipment by employing toxic micro-organisms and biological toxins requiring a level of containment and safety not now available within the Department of Defense". The request notes that the laboratory will be designed for working with "substantial volumes" of toxic biological aerosols.

Sasser, in a letter to Secretary of Defense Caspar Weinberger, asked that the full proposal be submitted as part of a supplementary budget request, which could be thoroughly examined and debated by the full Congress.

Stephen Budiansky

Tale of two cities

THE European Synchrotron Radiation Source (ESRS) just will not lie down. The decision (see *Nature* 1 November, p.3) to put the 5-GeV machine on a site at Grenoble in France, next to the high-flux neutron reactor of the Institut Laue-Langevin (ILL), is now being questioned. The site, critics say, is cramped and provides no room for expansion. But the French research minister, Hubert Curien, says that he made sure that Grenoble really did have "a few" possible places for the machine.

Meanwhile the city of Strasbourg, which also had plans to host the ESRS (and maybe even a site) is so upset with the decision that the Alsace Regional Council has threatened to boycott a visit next week by the French President, François Mitterrand.

The Alsatian feeling, it seems, is that M. Curien had promised ESRS to Strasbourg months ago — and in writing — and that he was forced to back down because of political pressure from above. Alsations say that the socialist party had long held Grenoble city council until they lost it in local elections a year ago, and that to have refused the city the synchrotron source would have been an electoral death-knell.

Curien, however, described such interpretations as "your responsibility", while pointing out that ILL technicians will give ESRS instrumentation a quick start, that the existing neutron and synchrotron light user community prefers Grenoble and that Britain (which has not yet offered cash) also preferred Grenoble. Grenoble, he claimed, was simply the better site.

Robert Walgate

US air pollution

Tall stacks out

Washington

ACTING under a court order, the Environmental Protection Agency (EPA) has issued new rules that would significantly restrict sulphur dioxide emissions from power plants. Although formally the rules address only an arcane debate over what values to use for stack height in computer models of power plant emissions, their effect will be to reduce SO₂ emissions by 800,000 tons to 2.8 million tons per year throughout the United States. The total annual industrial production of SO₂ in the United States is put at 24 million tons.

Because most of the 600 power plants affected by the rule are in the Ohio Valley, the rules may also make a dent in acid deposition in the northeastern United States and Canada, which is believed to originate largely in SO₂ emissions in the midwest. The Reagan administration has repeatedly argued that more studies of acid deposition are needed before regulations can be justified. EPA officials acknowledge that the stack-height rules are not the most cost-efficient way to reduce acid deposition; they may in fact be the most expensive pollution control rules issued by the Reagan administration. EPA estimates put the cost at \$900 to \$4,600 million in capital investment and \$300 to \$1,400 million in annual operating expenses.

A plan last year by EPA administrator William Ruckelshaus to attack the acid deposition problem by cutting back SO₂ emissions by some 4 million tons per year was turned down by the administration.

The dispute over stack height goes back to the early 1970s, when EPA decided that pollution controls on power plants would be determined case by case. Rather than imposing fixed emissions standards, EPA set ambient air standards; computer models are thus required to relate plant emissions to ambient air quality.

Many utilities found that the cheapest way to avoid exceeding ambient air standards around their power plants was to disperse pollutants through tall stacks. Some built stacks as tall as 1,200 feet. In 1977, Congress acted to close this loophole by requiring that, for the purposes of the computer models, stack height should be considered to be no more than that required for "good engineering practice". In effect, utilities were not to be given credit for the extra dispersion of pollutants that the tall stacks provided.

Since then, the dispute has raged over the definition of "good engineering practice". The new rules offer a much more restrictive definition than that originally proposed by EPA, which allowed credit for stacks tall enough to send smoke plumes over nearby mountains. Thus affected plants must switch to low-sulphur coal or install scrubbers to remove SO₂ from the stack.

Stephen Budiansky

British innovation

Research corporation relaunched

BRITAIN'S own academic entrepreneur, the National Research Development Corporation (NRDC), emerged more or less unscathed last week from the marriage into which it was forced in 1981 with the body called the National Enterprise Board, itself a relic of the 1964 Wilson government's decision to use large amounts of public money to reshape the pattern of British industry.

During the past year, the merged partnership, called the British Technology Group (BTG) since 1981, has been selling off its investments in large companies, and BTG's chairman, Mr Colin Barker, explained last week that disposal of equity stakes would continue, and that the outstanding interest in outside companies was already less than £50 million.

For the future, the group plans to spend up to £15 million a year in the next five years in support of technology transfer between British researchers in the public sector and general industry. Mr Barker insists that the group will function as a broker, not as a venture capitalist, for innovations arising in universities and government-supported laboratories.

BTG has already held a meeting with more than 30 vice-chancellors of British universities at which its new range of services was described. As well as the patenting and protection of innovations at its own expense, the group now promises to share its net revenue on all newly arising projects with those responsible for them, and will also collaborate in development projects on a fifty-fifty basis.

One curious aspect of what is in effect the re-launch of NRDC is that the government seems not yet to have decided when BTG will lose its exclusive right to the exploitation of inventions arising in the British public sector. This change was first announced by Mrs Margaret Thatcher a year ago. Nor, according to Mr Barker, has it been decided when there will be legislation to formalize the *de facto* merger, four years ago, of the two parts of BTG.

As a business, BTG itself is an enterprise of modest scale. NRDC, in the financial year to 31 March last, earned a surplus of £5.18 million (before tax) on an income of £16.87 million, mostly from income from licensed innovations. Income last year had fallen sharply from that in 1982-83 (from £27.38 million) but profit increased, from £2.33 million to £5.18 million.

Revenues from synthetic pyrethrin pesticides (developed at the Rothamsted Experimental Station) are now rising steadily, and BTG has a handful of other potential money-spinners on its books. Mr Barker argues that BTG will be both profitable and self-financing during the five years ahead, the span of time for which the government has approved its business plan.

The management seems also to recognize that it will in future have to compete with other sources of financial backing for innovations, and estimates that the City of London has raised rather more than £400 million for these purposes in the past five years. BTG hopes to intervene earlier in the process of technology transfer between the public sector and private industry, but said last week that participation in more mature ventures such as the biotechnology company Celltech and the Agricultural Genetics Company (AGC), while exceptional, would not be ruled out.

Whether BTG will be a more comfort-

able partner than NRDC for universities in the process of innovation remains to be seen. On the common academic complaint against its predecessor that many projects would be turned down only after long delay, Mr Steven Dollands, BTG marketing director, said he hoped the group would be able to improve its performance. To strengthen links with universities, it has recruited eight liaison officers who will scour the universities for ideas and also has a scheme whereby sums of up to £20,000 can be committed without formality to particularly attractive projects.

BTG's success may also be limited by the scale of its planned expenditure, now only a small part of funds flowing to universities from commercial sources. But BTG is adamant about its role. **John Maddox**

Hepatitis B

Taiwan plans mass vaccination

TAIWAN, with a population of 18 million, has embarked on a world first in vaccination: a 10-year mass vaccination campaign against hepatitis B, which is endemic in the country and in many cases results in liver cancer. The vaccine to be used is serum-derived and French — news which marks an enormous step forward for the French producers, Institut Pasteur Production (IPP), the industrial offshoot of the Institut Pasteur Fondation research laboratory in Paris, and an equal step back for IPP's competitors (notably Merck, Sharp and Dohme of New Jersey).

This year and next, Taiwan will be using the IPP vaccine to immunize, free of charge, all newborn babies who are at obvious risk from infection from their mothers. In 1986 the campaign will be extended to cover all 400,000 babies born each year in Taiwan, and will then gradually extend to older children and ultimately to adults. IPP has signed a contract to provide a million vaccine doses over two-and-a-half-years, and to transfer production know-how. Sanofi, the French drugs company which for the time being owns 51 per cent of IPP, will transfer production equipment at cost, and will also offer cost-price production training.

Moreover, Pierre Tiollais's group at the Institut Pasteur Fondation research laboratory has reached "production stage" with a genetically-engineered (second generation) vaccine produced in animal cells. Thus if the Tiollais vaccine works well, and receives health authority approval, the vast Taiwanese market for second-generation vaccines may also be tied up with France.

This French success has not been a matter of marketing or clever pricing, Sanofi claimed on Monday, but of the technical superiority of the IPP vaccine. The Merck serum-derived antigenic particles "lack some important polypeptides", according to Sanofi, and so the IPP vaccine is more effective. The missing portion is the polypeptide expressed by the

"pre-S" region of the hepatitis-B viral coat gene, removed in the Merck production process when the product is treated with pepsin to remove serum proteins (once feared as they might induce autoimmunity). The IPP process does without pepsin. But now the pre-S region appears to be important in generating full immunity to hepatitis B.

Moreover, the Institut Pasteur may also have stolen a march on the next generation of hepatitis-B vaccines, which will be genetically engineered. Here there are four main choices, according to Tiollais: to produce the antigen in yeast or in animal cells and in either case to produce the main hepatitis-B surface (HBs) antigen (which the Merck vaccine is left with) or HBs plus the pre-S polypeptide as well. Early attempts (as at Biogen, now licensed to Wellcome, and also at Merck, Sharp and Dohme) used yeast and produced HBs alone. But Tiollais and IPP now claim to be at production stage with an animal (Chinese hamster ovary) cell line expressing and secreting full HBs plus pre-S antigens assembled into particles almost identical to natural (human) hepatitis-B particles. It has not yet been proved, "but it is a working principle", says Tiollais, that such a vaccine will be more effective than other less complete products.

Merck, Sharp and Dohme meanwhile have licensed their serum-based technology to a Singapore company called Singapore Biotech, and are believed also to have transferred second-generation yeast technology to Singapore. The latter product, however, lacks pre-S, though moves are being made to introduce it.

Thus it seems that the Pasteur, with a foot in the Taiwan market with the serum vaccine, may have a chance also of taking the second-generation market, particularly as a part of the new deal with Taiwan involves the Pasteur "sharing knowledge" with the Taiwanese Development Centre for Biotechnology. **Robert Walgate**

Protecting experimental animals

SIR —Christine Stevens, president of the Animal Welfare League, presents (*Nature* 27 September, p.295) a convincing case that some animals are being abused in medical research. One can only sympathize with the animals and feel indignant at those persons who are involved. Her remedy, however, which is to have new laws enacted, when the problem according to her, is the failure of enforcement of laws already on the books, appears to be wide of the mark. If Mrs Stevens wishes to rally support for her concerns, there are steps she could take which would help breach the barrier between animal welfare groups and the public, including those engaged in animal research. These animal welfare groups should: (1) state unequivocally that they recognize the necessity for some types of animal experimentation and cite specific areas of research that they support; (2) advocate legislation to provide improved care of animals and better supervision of facilities only if the laws contain provisions to pay for the extra expenses involved. At present, the added costs come from the already constricted funds available for research. There is a widespread suspicion that the real aim of such legislation is to further reduce all animal research. It would be simple to disarm this suspicion if it is not warranted, and (3) actively support the availability of a small percentage of unclaimed stray animals for approved research in accredited and supervised laboratories.

Unless such simple steps are taken and acted upon with sincerity, the public can only conclude that the animal welfare groups do not approve of the development of procedures such as the coronary by-pass operation, advances in microsurgery and other attacks on problems which afflict human beings. The decision that animals should be used in research for the benefit of mankind was not made lightly. However, just as the decision that stray animals must be removed from the city streets and destroyed each year involved the searching of our consciences, there are also situations, such as the need for animal research, which cannot be avoided. It is in the power of Mrs Stevens and groups concerned with animal welfare to take the few steps suggested above to the betterment of animals as well as mankind.

DANIEL L. KLINE

Department of Physiology,
University of Cincinnati Medical Center,
Cincinnati, Ohio 45267, USA

SIR — David Britt, in his Commentary on ethical committees for animal experimentation (*Nature* 11 October, p.503), suggested that some major animal welfare groups in Britain could be relied on to participate in any initiatives to establish these committees. We at UFAW (Universities Federation for Animal Welfare) have for

several years favoured the establishment of local committees to look at the practical and ethical problems raised by proposed investigations involving animals. (See UFAW's evidence to the Select Committee on the Laboratory Animals Protection Bill (House of Lords), Vol. 2, 1980.) Some ethical functions may already be carried out by research planning committees where these exist. Whether an existing committee is used or a new one is set up, the terms of reference should include consideration of the following: (1) The likelihood of the project achieving its declared purpose. (2) The suitability of the type of animals proposed for use. (3) The proposed use of anaesthesia and analgesia. (4) Ways in which the use of living animals could be reduced or avoided. (5) The competence of the people available to carry out the proposed techniques. (6) The suitability of the facilities available for after-care. (7) Whether, if pain or suffering is likely, the result being sought justifies the use of living animals.

The government's recent White Paper on *Scientific Procedures on Living Animals* proposed a system for licensing projects which requires a sponsor to give an opinion on certain of these questions. We suggest that it would of great benefit to the sponsor as well as to the applicant and to the Home Office Inspector if proposed projects could be discussed at an early stage by a local review committee, before being submitted to the sponsor.

We agree with Dr Britt that the workload of such a committee could, if necessary, be reduced if it were required to investigate in depth only those projects likely to cause more than trivial pain or suffering. We also agree that no useful purpose would be served by the presence of a convinced opponent of animal research. Any lay representative should have an open mind on the subject. It would be up to the proposer to prove the worth of the project.

The meeting held by the Liverpool Animal Ethical Group in October, referred to by Dr Britt, was most successful. When the ideas put forward have been published, the initiative should then be taken up by the scientific community in general and by the universities in particular.

JENNY REMFRY

Universities Federation
for Animal Welfare,
8 Hamilton Close, South Mimms,
Potters Bar, Herts EN6 3QD, UK

Seascale and cancer

SIR — Andrew Pomiankowski (*Nature* 13 September, p.100) draws attention to the work of Craft, Openshaw and Birch¹. This epidemiological study examines the incidence of diagnosed lymphoid malignancy in children under the age of 15 within the 675 census wards of the northern

region, during the period 1968 to 1982. Pomiankowski points out that, in this study, the wards are ranked in order of descending Poisson probability, and that the Seascale ward is ranked first (with a probability of occurring through chance of about 1 in 7,500), whereas in the report of Sir Douglas Black² the wards are ranked in order of descending incidence rate, in which the Seascale ward is ranked third (with a rate of 9.73 per 1,000 compared with a regional rate of 0.61 per 1,000). These statistics are based upon 4 cases in a sample of 411 children.

I agree that ranking according to descending Poisson probability is the more statistically appropriate procedure for indicating the significance of an incidence rate in a ward. However, I would question whether the inference that the Seascale ward is thus "unique" is justified. Craft *et al.* point out that: "For other varieties of childhood cancer there is a similar spread of 'highly ranked', but different, wards throughout the region". Indeed, according to Table 2.18 of the Black Report, which gives the ward incidence rates of all childhood malignancies, there is a ward (not Seascale) with an incidence rate having a probability of occurring by chance of about 1 in 3,000 if sampled from a Poisson distribution. Thus the pattern of distribution of childhood lymphoid malignancies in the region would not appear to be markedly different from that of other malignancies.

The results of the analysis of Craft *et al.* also show that Seascale is the only ward in western Cumbria with a childhood lymphoid malignancy rate significantly higher (at the 0.05 level) than the regional mean rate. If a causal relationship is being suggested between excess childhood lymphoid malignancies and an increased radiation exposure of individuals due to the operations of the Sellafield site, then it might be expected that the environmental distributions of discharged radionuclides³ would give rise to higher leukaemia rates in more than just one local ward of the area.

However, the other data in the Black Report do indicate that there are complicating factors in the accurate interpretation of leukaemia incidence, such as length of analysis period, the age range of relevance, and time resident in the district. As pointed out in the report, further epidemiological and radiological studies are clearly needed. Meanwhile, caution is required in drawing conclusions about the role of radiation in the aetiology of childhood lymphoid malignancies in West Cumbria.

RICHARD WAKEFORD

6 Bourne Street,
Wilmslow, Cheshire SK9 5HD, UK

1. Craft, A.W., Openshaw, S. & Birch, J. *Lancet* II, 96-97 (1984).
2. Report of the Independent Advisory Group Investigation of the Possible Increased Incidence of Cancer in West Cumbria (HMSO, London, 1984).
3. Cambray, R.S. *et al.* *Studies of Environmental Radioactivity in Cumbria* AERE Harwell, UKAEA, (1980-82).

Geological Survey

SIR — With regard to your succinct appraisal of the Natural Environmental Research Council (NERC) and, more particularly, of the British Geological Survey (BGS) and the problems of its basic mapping programme (11 October, p.499), it is important to realize two things. First, the geological survey of the country is a fundamental need. It forms the starting point both of countless "academic" researches and of even more investigations with a practical purpose — geotechnical, resource-seeking, planning or whatever. It should be supported on a stable basis, irrespective of short-term advantages or requirements, and should certainly not be at the mercy of a change of government, as it was when the "Rothschild principle" was brought in in 1973.

Second, it needs long-term support because it is never finished. Ideas and needs are both constantly changing. An area that was mapped, even to the highest standards, 40 or 50 years ago may need doing again, because progress in geology will result in things being found which were previously unremarked, and because new kinds of information must be gathered to meet new demands. Hence there must be regular revision and long-term planning.

You quote a BGS figure of "at least 25 per cent" of the maps as obsolete. This appears to represent maps not available at all plus those listed as provisional editions, that is, based on very old mapping. It is a gross underestimate. Many sheets, not classed as provisional, are based on mapping in the early years of this century, at the latest. To take an example at random, Sheet 266, listed in 1983 as published 1974 (in fact the date when it was transferred from the old one-inch to the 1:50,000 scale), was surveyed 1888–95, partially revised 1923–24. The oldest work on it is now 96 years old. Many other sheets have a similar history.

When we consider the accompanying sheet memoirs the situation is scandalous. They are essential because, obviously, not all the information gathered by the surveyors can be shown on the map. Yet the list of June 1983 shows only 60 memoirs as available, that is for about 22 per cent of the available maps. Even adding the 50 memoirs said in your article to be in preparation only raises this figure to about 40 per cent. The survey's inability to get its memoirs written and published is legendary. Some of those that can be bought were published only several decades after field work was completed.

There should be, urgently, public discussion of the survey's long-term plans, and a programme drawn up which will allow real progress, that is, maps and memoirs being produced faster than they are becoming obsolete. Funds for this must be earmarked and must not be at risk from the numerous other activities of the survey. Then, an organization needs to be devised

which will get work completed and published in reasonable time. Many of the missing memoirs were in people's heads. They just never got written.

The Geological Survey's 150th anniversary will occur in 1985. It could receive no better birthday present than an assured and planned future.

D.T. DONOVAN

*Department of Geology,
University College London,
Gower Street, London WC1E 6BT, UK*

Unfair to SERC

SIR — *Nature* has a reputation for sharp but informed criticism of the activities of the UK research councils. However, your leading article of 4 October (p.397) was doubly unfair. After discussing the "inflexibility" of the research councils you criticize the Science and Engineering Research Council for taking on burdens of "strategic" research and then refer disparagingly to funds earmarked for "trendy" subjects: you cannot have it both ways. In my experience, it is the refusal of the research councils to follow every "fashionable" topic that leads to accusations of "inflexibility". These accusations often come from people who think their own line of research is being starved — anything else, particularly long-term studies that need continuity, is to them "inflexibility".

Scientists who have the fortune to work in certain organizations will know that it is perfectly possible to combine "strategic" research and flexibility. A long-term series can be continued, but at the same time the scientists involved can alter the course of part of their studies in response to results of their own and their colleagues' researches. It can be bad for a scientist to pursue the same narrow line all the time, but equally bad to take up "trendy" subjects for the sake of newness.

The situation we face now is not unique. Arguments about the relative values of short-term contracts, once called "industrial" research, compared with open-ended "spontaneous" research have raged for at least a hundred years. In my own discipline, as long ago as 1920, in an unpublished but influential report, Sir Ray Lankester pointed out the incalculable advantages of "spontaneous" research in which the solution of one problem leads to the formulation of more problems to investigate, problems that were not apparent at the start of the programme. Everyone will have examples from his or her own discipline. This sort of research is now condemned by critics who favour short-term contracts. Many of these short-term contracts fall into the "trendy" category, others are "safe" topics thought up by committees: what major scientific advances have accrued since the inception of the short-term "customer-contractor" principle for government-financed research?

Nature should avoid this sort of general attack on British research and come out more in favour of "spontaneous" research as the *fons et origo* of scientific advance.

A.J. SOUTHWARD

*108 Durnford Street,
Plymouth PL1 3QP, UK*

Carbon dioxide

SIR — In her article in *Nature* (4 October, p.401), Anna Lubinska erroneously assumes that carbon dioxide is a harmless gas, whereas it is in fact corrosive. Millions of tons of the gas enter the atmosphere every day in the products of combustion from fossil fuels such as coal, oil products, natural gas and other sources.

The carbon dioxide dissolves in rain water to form corrosive carbonic acid that falls on soil, streams and lakes, but also on calcareous materials such as limestone rocks, concrete products and marble. The acid dissolves these materials to form soluble calcium bicarbonate. Copper and bronze are attacked to form cupric carbonate — commonly called verdigris.

The other chemicals listed occur at levels which are infinitesimal compared with the quantity of carbon dioxide in the products of combustion of petrol. The modified list reads:

	grams per mile
Hydrocarbons	0.41
Carbon monoxide	3.4
Nitrogen oxides	1.0
Carbon dioxide	320.0

The quantity of carbon dioxide is increased by 5.3 grams when the carbon monoxide is rapidly oxidized in the atmosphere.

P.P. JONES

*Cercol Laboratories,
Conways Drive, Parkstone,
Poole, Dorset BH14 0PL, UK*

Chemical waves

SIR — It cannot have escaped the attention of some of your readers that the photographic representations of spiral chemical waves (*Nature* 18 October, pp.611–615) bear an uncanny resemblance to the prehistoric carvings of spirals on stones. In



Spiral carving, Newgrange passage grave, Co. Meath, Ireland (right) and a monolayer of slime mould cells (left)?

spite of numerous studies, these ancient carvings have so far defied interpretation. Is it conceivable that our ancestors understood chemical waves?

RICHARD A. BATCHELOR

*12 Main Street, Strathkinness,
St Andrews, Fife, UK*

Cataclysmic calderas catalogued

The centenary of the volcanic eruption of the island of Krakatoa has been an occasion for listing other such phenomena. There are (and have been) a lot of them.

ONLY volcanologists would think it proper, not macabre, to celebrate the centenary of a natural disaster, the eruption on Krakatoa in 1883, with a symposium. The event did, after all, put their subject on the map — but removed Krakatoa itself. The papers gathered for the fall meeting of the American Geophysical Union (AGU) last year, have now been published in an extra-thick issue of *J. Geophys. Res. B* (30 September, 89, 8219-8842; 1984). The volume is, as it turns out, a kind of Baedeker of the known roughly circular pits called calderas, relics of "the most catastrophic geologic (*sic*) events that have occurred on the Earth's surface..."

Part of the trouble with extinct calderas is that they are often inconspicuous. Recognizing a nearly circular depression of 100 metres or so which may be 10 km or more across is easier from satellites than from the ground, especially when once-sharp edges have been eroded away. This is how R.L. Wunderman and W.I. Rose are able to describe for the first time (89, 8525) a still-active caldera just over 10 km south of Guatemala City.

Other calderas are more conspicuous, however. Carter Lake in Oregon is one, while California, Nevada, Utah and New Mexico are peppered with them, as is the East African Rift Valley. One of the neatest is Deception Island, a ring of barren rock in the South Atlantic. Olympus Mons on Mars (60 km across), splendidly photographed by the Viking orbiting spacecraft, may by now be one of the most familiar.

So this collection of articles is in part a kind of guidebook to known major calderas. The western United States are well worked over, but L.A. Morgan *et al.* (89, 8665) are able to make a case for three calderas in Idaho, which may have been the source of 1,500 km³ of ash (now tuff) about 6.5 million years ago.

Further south, just across the Mexican border with California, D.L. Nielson and J.B. Hulen (89, 8695) have been able to use data gathered during geothermal drilling into the familiar now-domed Pliocene Valles caldera to construct a three-dimensional geological section through the structure. The 1,000 metres or more of tuff is interspersed with sandstone of irregular thickness — there was once a lake — and the section shows how the tuff has slumped along nearly vertical planes from time to time. But although the Valles caldera has now been converted into a domed structure by some 900 metres of recent uplift, even

the deepest borehole (3,242 metres) has failed to intersect the rock responsible.

The richest caldera country in the world is the remarkable 65,000 km² of New Mexico, west of the Rio Grande, that seems to have suffered 20 million years of volcanism beginning in the late Eocene, 40 million years ago. W. E. Elston (89, 8733) lists 28 calderas so far recognized in the region, many of which have been put conveniently on display by erosion and faulting during the past 20 million years.

The sheer scale of this volcanism has evidently been a serious impediment to its interpretation. Elston writes wistfully of the early puzzle of the Black Range, 50 km west of the township of Truth or Consequences, a dome-shaped structure 55 km long, 2 km wide and 2 km high, now recognized as a 34-million-year caldera, whose central depression has been turned inside-out by the resurgence of magma. Can anyone now doubt that the Colorado Plateau, through which the Grand Canyon cuts, is a magmatic structure of the same kind?

The AGU symposium seems to have been generally agreed on the mechanism by which calderas are formed. A subterranean magma chamber contaminated with volatile components, water at the very least, breaks through the overburden, explosively to begin with. The volume of material ejected in one or more eruptions may be huge, but the characteristic of a caldera is the eventual collapse of the surface into the magma chamber beneath, at which stage there may be further eruptions through fissures or from the ring fracture around the subsiding plate.

Elston's catalogue of New Mexico includes all possible outcomes — calderas filled to the brim and beyond with the ash from post-collapse eruptions, others where resurgent doming has preceded roof collapse, and then vice versa. Individual calderas have been responsible for 2,000 km³ of volcanic effluent. Halfway through the mid-Eocene volcanic episode, at least some magma chambers appear to have reached to within 2 or 3 km of the surface.

A few simple lessons stand out from the handful of attempts to account for these happenings. Simulations of the initial eruption of an over-pressurized magma chamber by calculations of what happens in a shock-tube (where the diaphragm between regions of high and low pressure gas is broken; K.H. Wohletz *et al.*, 89, 8269) show that it is reasonable that the first ash-laden plume should reach to the strato-

sphere and that horizontal winds (laden with ash) may travel at 200 m s⁻¹ or so at distances of 10 km or more from the vent. The mechanism by which caldera floors collapse is less clear, especially because many ring-shaped sunken floors seem to have fallen through downward-tapering apertures. Resurgence may be a precondition, and may itself arise from non-equilibrium processes in the magma chamber, compensation of over-expulsion of volcanic material or slow dissolution of volatiles from magma rendered metastable by an eruption (B.D. Marsh, 89, 8245).

What stands out a mile, however, is that nobody has a clear idea of how the magma chambers have appeared within a few kilometres of the Earth's surface. It would be comforting if it turns out that magma chambers are not like subterranean caverns filled with molten mantle material from below but places within the crust in which ordinary material behaves so differently from its immediate surroundings that it seems qualitatively a liquid, not a solid.

The Long Valley caldera (with Mammoth Lake within its rim) due east of San Francisco may have the answer. The floor collapsed 720,000 years ago, but is now rising, at some places by 5 cm a year. Two underground magma chambers within 6 km of the surface have been mapped seismically (C.O. Sanders, 89, 8287). Measurements on the spot do not indicate whether the outcome will be stasis or another eruption — but 5 cm a year is a lot.

In all this, the ubiquity of calderas is the big surprise. But most known calderas are young, suggesting that most of them do not leave permanent traces. Krakatoa, after all, has all but vanished.

So, too, has Minoan civilization. The spell-binding paper in this symposium is the reconstruction of the history of the caldera on Thera (Santorini) in the Aegean, whose eruption about 3,400 years ago put paid for a time to nearby Crete (G. Heiken and F. McCoy 89, 8441). The eruption appears to have come from a volcano on dry land, to the east of a submarine caldera active during the previous million years. There were four distinct phases, perhaps separated by several years, but the first eruption that put Crete in trouble was preceded by minor ash-fall that would have given the people of Thera a chance to get away, which explains why no people have been found among the ashes. And volcanic activity continues, at the join of two submarine calderas.

John Maddox

Bacterial physiology

New kingdom for nitrogen fixation

from John R. Postgate

THE report¹ by Murray and Zinder on page 284 of nitrogen fixation by a strain of methane-producing bacteria may seem to be just another addition to the list of nitrogen-fixing bacteria. In fact, it is something of a landmark in bacterial physiology, as it provides the first example of diazotrophy — the ability to 'fix' atmospheric nitrogen by converting it to a metabolically available form — outside the kingdom of 'regular' bacteria.

The methane-producing bacteria or methanogens, which form 'marsh gas' from decaying vegetation and which inhabit sewage sludge, the rumens of ruminants and the inner depths of aqueous sediments, are a difficult group of bacteria to study. They are very strict anaerobes, completely inactivated by exposure to air; they require complex media, with fixed nitrogen plus reducing agents, to grow in the laboratory; and they are difficult to obtain in pure culture because they often cohabit with other anaerobes, particularly sulphate-reducing bacteria². Indeed, a report thirty years ago³ of nitrogen fixation by the methanogen *Methanobacillus omelianskii*, soundly based on uptake of ¹⁵N₂ by the bacteria, collapsed when the 'species' was shown to be a consortium of two different bacteria⁴. There was no evidence to show whether the methanogen rather than its consort was responsible for the diazotrophy.

There the matter rested for many years; though new diazotrophs were reported among other bacterial groups, the sheer awkwardness of the methanogens, and perhaps the relative dullness of adding yet another diazotroph to the list, seems to have inhibited further examination of methanogens for diazotrophy.

The matter became interesting again around 1980, when the methanogens, together with certain halobacteria and sulphur-metabolizing bacteria, were shown to belong to an entirely new group of living things: the archaeobacteria⁵. Early scepticism has been dispelled, and we now know that the archaeobacteria comprise a separate kingdom, which differs from the other two kingdoms, eubacteria (regular bacteria) and eukaryotes (multicellular organisms), more than those kingdoms differ from each other. The separateness of the archaeobacteria was first established from 'fingerprints' of their ribosomal RNA and is reflected in a variety of biochemical properties unique to the group⁶.

Murray and Zinder's report of diazotrophy in *Methanosarcina barkeri* is the first definitive demonstration that nitrogen fixation occurs among the archaeobacteria. It is definitive because it includes two important controls. The first is that uptake of

¹⁵N₂ is shown to accompany growth, so there is no possibility of simulated diazotrophy arising from growth on nitrogenous impurities in the gases or reagents, a problem which has led to many false reports of diazotrophy⁷. The second and perhaps more important control is that the authors have shown to the best of their ability that their culture is pure and, in particular, have excluded sulphate-reducing bacteria, which are known to be diazotrophic⁷.

Genetic and biochemical studies of nitrogen fixation among the eubacteria are converging to suggest that the genetic determinants for diazotrophy, the *nif* genes, are highly conserved and that their products, nitrogenase in particular, are much the same whatever their genetic background. The *nif* genes are fairly readily transferred among bacteria by experimental means — wholly new diazotrophs have been generated in this way — and there are compelling reasons to believe that they represent a genetic linkage group that has spread naturally among the eubacteria⁸. It has not spread to the eukaryotes — as far as we know — and all eukaryotic nitrogen-fixing systems are in fact biocoenoses involving mutualisms with diazotrophic eubacteria, the legume-rhizobium symbioses being the most familiar example. It is fascinating, then, that diazotrophy should now be found among the archaeobacteria.

Did it spread to the archaeobacteria from the eubacteria or, rather, originate in the archaeobacteria and spread to the eubacteria? Could it be so ancient a property that it originated before the archaeobacteria and eubacteria diverged, very early in the history of life on Earth⁶? Or could diazotrophy have originated separately in the two kingdoms? A study of *nif* gene organization and the biochemical nature of nitrogenase in a methanogen ought to be most instructive, for the possibility exists of

finding nitrogenase enzymes that differ substantially from those of eubacteria.

A hint that the archaeobacterial nitrogenase system is different comes from the report⁹ on page 286 of this issue by Belay, Sparling and Daniels that a presumptive diazotrophic methanogen, *Methanococcus thermolithotrophicus*, functions at 64°C. The temperature maximum for diazotrophy in eubacteria is generally 35–40°C (although the cyanobacterium *Mastigocladus* can just about manage at 60°C). The reason appears to be that, although nitrogenase itself is reasonably thermostable, the genetic activator for *nif* expression (the *nifA* product) is thermolabile⁸. Belay, Sparling and Daniels's finding could imply that archaeobacterial *nif* systems can be unusually temperature-tolerant, so it is regrettable that their report omitted the meticulous controls used by Murray and Zinder. Thus, absence of sulphate-reducing (or other) eubacterial thermophiles was apparently not checked and, secondly, error due to inadvertent introduction of fixed nitrogen into the experimental system was not rigidly excluded by using ¹⁵N₂. Their success in obtaining acetylene reduction, a test avoided by Murray and Zinder in this context because of its obliquity, tends to support their claim for thermophilic diazotrophy. But some doubt must linger about whether the organism responsible is an archaeobacterium and not some diazotrophic commensal eubacterium. □

1. Murray, P.A. & Zinder, S.H. *Nature* **312**, 284 (1984).
2. Stadtman, T. & Barker, H.A. *J. Bact.* **61**, 67 (1951).
3. Pine, M.J. & Barker, H.A. *J. Bact.* **68**, 589 (1954).
4. Bryant, M.P., Wolin, E.A., Wolin, M.J. & Wolfe, R.S. *Arch. Microbiol.* **59**, 20 (1967).
5. Fox, G.E. *et al.* *Science* **209**, 457 (1980).
6. Stackebrandt, E. & Woese, C.R. in *Molecular and Cellular Aspects of Microbial Evolution*, Symp. Soc. gen. Microbiol. **32** (eds Carlile, M.J., Collins, J.F. & Moseley, B.E. B.) 1 (Cambridge U. Press, 1981).
7. Postgate, J.R. in *Current Perspectives in Nitrogen Fixation* (eds Gibson, A.H. & Newton, W.E.) 217 (Aust. Acad. Sci., Canberra, 1981).
8. Postgate, J.R. *The Fundamentals of Nitrogen Fixation* (Cambridge U. Press, 1982).
9. Belay, N., Sparling, R. & Daniels, L. *Nature* **312**, 286 (1984).

John R. Postgate is Director of the Agricultural and Food Research Council Unit of Nitrogen Fixation, University of Sussex, Brighton BN1 9RQ, UK.

Atmospheric chemistry

Ozone fears revisited

from H.I. Schiff

THE threat to the Earth's ozone layer promises to be a saga with as many sequels as *Star Wars*. Just as we have become assured that this protective shield is more robust than had been feared, M.J. Prather, M.B. McElroy and S.C. Wofsey warn that the Empire might strike back. In a paper on page 227 of this issue, they show that large decreases in ozone can occur if the emission of chlorine compounds into the stratosphere is allowed to increase by a factor of 5 or more.

The ozone layer is located in the stratosphere, between about 10–50 km above the Earth's surface. Although more than 95% of atmospheric ozone is contained in the stratosphere, its concentration never amounts to more than a few parts per million of air. And yet its presence is essential for life on this planet, since it prevents biologically-damaging solar ultraviolet radiation from reaching the surface. It also plays an important role in determining the climate of the planet.

The small concentration of stratospheric ozone represents a balance between photochemical processes which produce it and others which destroy it. More than 50% of the destructive processes are controlled by catalytic chains involving compounds containing nitrogen (NO_x), hydrogen (HO_x) and chlorine (ClO_x). The recognition that these compounds are present in concentrations thousands of times smaller than ozone itself, is the basis for the concern that human activities might increase their concentrations and thereby lower the steady-state concentration of ozone. The first of these potential activities to be identified was the operation of fleets of supersonic aircraft, which could add considerable quantities of NO_x to the stratosphere. When such fleets failed to materialize, the focus of attention shifted to chlorofluorocarbons (CFCs) as a more imminent threat.

Computer models were used to describe the current atmosphere and to predict the effects of various scenarios of future emissions of CFCs. These models use rate coefficients, measured in the laboratory, of the chemical reactions believed to occur in the atmosphere, and highly parameterized descriptions of air motions. The validity of the models can be checked by measurements of the current concentrations of a number of atmospheric constituents.

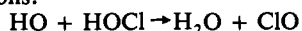
The predictions of these models for ozone depletion as a result of continuing emissions of CFCs have appeared in a number of reviews²⁻⁶ and have vacillated widely over the past 10 years. For example, take the US Academy of Sciences prediction in response to Congress's requirement for a biennial report on the health of the ozone layer. Assuming continual emissions of CFCs at their present rates, the Academy predicted in 1976 that the most probable reduction in the ozone level would eventually (in about 100 years) reach 6–7.5%. The figure was raised to 16.5% in 1979, lowered to 5% in 1982 and this year is 2–4%. Moreover, if emissions of other atmospheric gases, such as CH_4 , CO_2 and tropospheric NO also continue to increase, ozone depletion could be vanishingly small.

The main reason for these vacillations is that the stratosphere is a very promiscuous place — a great deal of coupling occurs between the NO_x , HO_x and ClO_x families. Current models incorporate about 200 chemical reactions between some 50 members of these families. Over the past 10 years a number of new reactions have been identified and new laboratory measurements have led to significant changes in the rate coefficients of many of the important reactions.

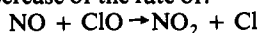
Prather *et al.* claim that this system of reactions could become dramatically non-linear if the amount of Cl compounds in the stratosphere were to exceed the amount of nitrogen compounds. This could occur if the emission rates of CFCs were not

constant, but were allowed to increase exponentially. Under those conditions a 10% ozone depletion could occur within 50 years for a 5% per year increase in CFC emissions, or within 75 years for a 3% per year increase. This prediction is similar to those made by Wuebbles⁷ and in the latest Academy report⁵.

The reasons for non-linearity have been analysed by Prather *et al.* Where CFC emission is constant, the ClO_x compounds are most effective in destroying ozone above 30 km, whereas the NO_x and HO_x chains dominate in the lower stratosphere. But, with higher Cl content, the ClO_x catalytic chains can be active at much lower altitudes. The increased chlorine ties up most of the NO_x in the form of non-reactive ClONO_2 . This, in turn, increases the amount of HO because NO_x reactions are the major sinks for HO radicals. The decrease in NO_x and increase in HO produce a more than linear increase in ClO , mainly because of increased rate of the reactions:



and a decrease of the rate of:



When chlorine compounds exceed the nitrogen compounds, ClO increases three times faster than total chlorine and becomes the most effective ozone destroyer down to altitudes of 20 km.

The current chlorine content of the stratosphere is 3 parts per 10^9 by volume (mixing ratio 3×10^{-9}) and would reach the NO_x concentration of 13 parts per 10^9 in 50 years if CFC emissions were to increase at a rate of 5 per cent per year.

Prather *et al.* do not discuss the likelihood of such CFC increases. The production rates of the two major CFCs, CFCl_3 and CF_2Cl_2 , have not changed significantly over the past six years⁸. During that period, the production of these compounds for use as aerosol propellants has been reduced by controls in several countries from 56% to 34% of the total propellant production (at least by those companies that have reported figures). On the other hand, the non-aerosol use of these compounds has increased by 4.0% per year over the period 1976–1982 in the same companies, and by as much as 6.5% per year worldwide⁹. Other chlorine containing compounds such as CCl_4 and CH_3CF_3 make smaller but significant contributions to stratospheric chlorine. Projections for the future emissions of all the human-produced chlorocarbons is extremely speculative and will depend on both economic and social factors. An increase of 3–5% per year into the next century is unlikely but not beyond the realm of possibility.

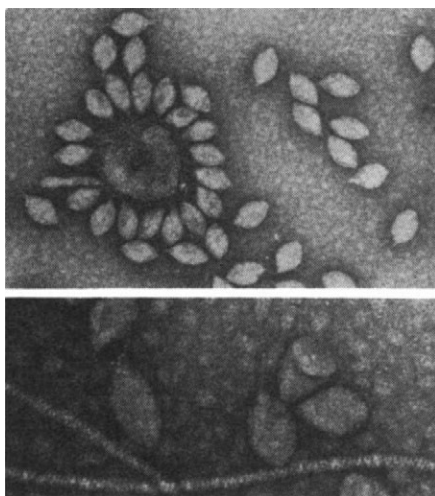
The authors have also considered the effects of increases in emissions of some other gases. Atmospheric concentrations of both CH_4 and N_2O have been observed to be increasing. Continuation of these trends would delay the time when the dramatic effect of CFCs would occur. On

the other hand, increased releases of bromine compounds, used largely as fire extinguishers and fumigants, would lead to additional ozone depletion. However, increased CO_2 emissions, not considered by the authors, which raise the tropospheric temperature but lower the stratospheric temperature, would increase the rate of ozone production and decrease the rate of its destruction.

For their calculations the authors have used a one-dimensional model which treats air motions only in the vertical direction, and not in latitudinal or longitudinal directions. Horizontal motions are particularly important in the lower stratosphere where the effects discussed by these authors occur. Nevertheless the potential effects are sufficiently serious to warrant careful consideration. Fortunately there is still time to monitor the increase in stratospheric chlorine and bromine before effects occur. □

1. Prather, M.J., McElroy, M.B. & Wofsy, S.C. *Nature* **312**, 227 (1984).
2. National Research Council *Halocarbons: Effects on Stratospheric Ozone* (National Academy of Sciences, Washington, D.C. 1976).
3. National Research Council, *Stratospheric Ozone Depletion by Halocarbons: Chemistry and Transport* (National Academy of Sciences, Washington, D.C. 1979).
4. National Research Council, *Causes and Effects of Stratospheric Ozone Reductions: An Update* (National Academy of Sciences, Washington, D.C., 1982).
5. National Research Council, *Causes and Effects of Changes in Stratospheric Ozone: An Update* (National Academy of Sciences, Washington, D.C., 1984).
6. WMO Global Ozone Research and Monitoring Project Rep. No. 11: *The Stratosphere 1981: Theory and Measurements* (eds. Hudson, R.D. *et al.*) (World Meteorological Organization, Geneva, 1982).
7. Wuebbles, D.J. *J. geophys. Res.* **88**, 1433 (1983).
8. *World Production and Release of Chlorofluorocarbons 11 and 12 through 1981*, Report FFP 83-F (Chemical Manufacturers Assoc., Washington, D.C., 1982).
9. Wuebbles, D.J., MacCracken, M.C. & Luther F.M., *A Proposed Reference Set of Scenarios for Radiatively Active Constituents*. Report No. DOE/NBB-0066, (Department of Energy, Washington, D.C. 1984).

H.I. Schiff is in the Department of Chemistry, York University, Ontario, Canada M3J 1P3.



Viral particles produced by the archaeobacterium, *Sulfolobus acidocaldarius*. The 100×60 nm particles, many of which can be seen attached to bacterial remains (above), are generally lemon-shaped but, perhaps when halved, occasionally appear bullet-shaped (below). (From Martin, A. *et al.* *EMBO J.* **3**, 2165; 1984.)

Reproductive strategy

Theory and practice of extramarital sex

from Jared Diamond

In many animal species there is a lasting pair bond between a male and a female, both of whom devote much parental care to their joint offspring. Under what circumstances should such shared parental care coexist with extramarital sex, a situation that sociobiologists euphemistically term 'mixed reproductive strategy'? A recent study by Fitch and Shugart (*Am. Nat.* **124**, 116; 1984) provides some answers.

On first reflection, one might think that males would always benefit and could never do themselves harm by pursuing a mixed strategy, because males make a much lower investment in each gamete than do females. Since the theoretical studies of Trivers (in *Sexual Selection and the Descent of Man*, ed. Campbell, B., 136, Aldine, Chicago; 1972), however, other factors have been appreciated. For example, extramarital sex may decrease rather than increase one's fitness if one's mate deserts as a result. A male pursuing a mixed strategy must have confidence of paternity, that is he must be assured that, while he is seeking another female, his mate is not seizing the chance to copulate with another mixed-strategy male. Perhaps the most fundamental theoretical problem is that mixed strategies cannot be considered from the standpoint of the male alone: males pursuing mixed strategies are doomed to failure in the absence of similarly inclined females. What could induce a female to pursue a mixed strategy?

Several field studies have reported evidence for mixed strategies in birds, but they have usually been based on unmarked individuals, so that it has been unclear whether participants in the supposed extramarital sex were paired, unpaired, or engaged in courting (which could lead to formation of a firm bond). That methodological problem has now been overcome by Fitch and Shugart, who marked individual herring gulls on Gull Island, Lake Michigan, and drew the following conclusions from lengthy observations.

Of the mated males studied, 35 per cent were seen to copulate with a female other than their mate. Of the mated females studied, none was seen to copulate with a male other than her mate, despite ample opportunity to do so. (For 14 per cent of the time during the fertile period, mated females were alone at the nest and a male was present on a neighbouring territory. The females did not approach that male, however, and rebuffed him if solicited). By contrast, unmated females usually did copulate if solicited.

Mated males spent more time in their territory during the period when their mate

was fertile than when she was not fertile. When another bird entered the territory, the male almost always chased out the intruder. These two observations suggest that males seek confidence of paternity by guarding fertile mates, although guarding is not the sole answer as fertile females were left unguarded for 20 per cent of the time.

Thus, both male and female herring gulls

pursue a mixed reproductive strategy, but males pursue both components simultaneously, while females do so sequentially. Males are accepted for extramarital sex by unmated females but not by mated females. Fitch and Shugart speculate that males induce their mates to remain faithful by diligently feeding them and by copulating frequently when the female is receptive. Ironically, none of the extramarital copulations witnessed by Fitch and Shugart resulted in an egg being laid. Thus the adaptive benefits of extramarital sex remain mysterious. □

Jared Diamond is Professor of Physiology at University of California Medical School, Los Angeles, California 90024, USA.

Cell biology

Microtubule catastrophe

from J. Richard McIntosh

MICROTUBULES are an important part of the system of cytoplasmic fibres that defines the shape of a cell and helps it to move around. Observations of these fibers with several kinds of microscopes have demonstrated that cells modulate the organization of their microtubules during the cell cycle. For example, the microtubules that are arrayed in the cytoplasm of the cell in interphase disassemble into tubulin subunits as the mitotic spindle forms. Two papers by T. Mitchison and M. Kirschner in this issue^{1,2} cast the lability of microtubules in a new light. Under certain conditions *in vitro*, a population of microtubules will behave in two ways at once: some polymers elongate while others catastrophically disassemble and disappear.

It is widely believed that the positions of microtubules in cells are defined by special 'organizing centers' composed of macromolecules other than tubulin. A cell modulates the activity of these centres to define the number of microtubules that form. The centrosome is probably the most important microtubule organizer in an animal cell. This organelle was recognized years ago as a regulator of cellular morphogenesis, but modern work on centrosomes has been sparse, in part because they are difficult to isolate for detailed study. On page 232 Mitchison and Kirschner present a method for preparing active centrosomes from cultured mammalian cells and describe some of the properties of the microtubule arrays they initiate. This study confirms earlier work on centrosome activity by showing that the number of microtubules initiated depends on the concentration of tubulin used for assembly³⁻⁵. If the subunit concentration exceeds a critical value, microtubules will form, as expected from the 'condensation-polymerization' model of Oosawa and Kasai⁶ but, unlike most phase transitions, the number of polymers formed increases rather slowly with

increasing concentrations of subunits, saturating at roughly five times the critical concentration.

Mitchison and Kirschner go on to identify properties of their system that are incompatible with the formulation by Oosawa and Kasai. The variation in microtubule number with the concentration of soluble tubulin is not simply a consequence of the amount of tubulin present at the time of polymer initiation. When microtubules initiated from a centrosome at high tubulin concentration (lots of microtubules per organizing centre) are diluted into an intermediate concentration of tubulin, the number of microtubules per centrosome decreases to that appropriate for the intermediate concentration. Even more surprising, while some microtubules are disappearing, the ones that are left get longer.

One might imagine that this behaviour reflects some unusual property of centrosome-initiated microtubules, but in their second paper, Mitchison and Kirschner show analogous phenomena for microtubules formed by spontaneous initiation in solution. Again, the distribution of lengths in a preparation of microtubules changes upon dilution, with some fibres elongating, while others disappear. At polymerization steady state, a constant polymer mass is maintained by the slow elongation of ever fewer microtubules. This view of the steady state is completely different from the one presented by Margolis and Wilson⁷, which holds that subunits add to one end of each polymer and fall off from the other, a process often called treadmilling. It is too soon to evaluate these two views of the steady state, but the distinctions between their predictions are very clear. Discrimination by experiment should soon be possible.

To interpret their data, Mitchison and Kirschner use a model based on studies of GTP binding to tubulin⁸⁻¹¹. GTP that

binds to active tubulin is hydrolysed to GDP in association with tubulin assembly, but not necessarily at the time of subunit addition. The delay in hydrolysis results in a cap of GTP-tubulin whose length depends on the rate of polymerization. This suggests that polymer growth should normally be a balance between the rates of addition and falling off of GTP-tubulin, but if GTP hydrolysis should catch up with subunit addition, the situation would change. The GTP-tubulin cap would then be converted to a GDP-tubulin end, and new polymerization reactions would become relevant. Two lines of evidence show that the rate at which GDP-tubulin falls off a microtubule exceeds the off-rate for GTP binding to tubulin by two-to-three orders of magnitude, leading to the rapid disassembly of any microtubule with an exposed GDP-tubulin end.

Mitchison and Kirschner postulate that the balance between GTP-tubulin caps and GDP-tubulin ends accounts for the microtubule length redistributions they have observed. At low tubulin concentrations, the GTP-tubulin cap is short enough that statistical variation gives rise to some microtubules with an exposed GDP-tubulin end. The rapid off-rate of this state will lead to a catastrophic disassembly of those microtubules, while the polymers that retain their GTP-tubulin cap will continue to elongate. For centrosome-initiated assembly, the disappearance of a microtubule exposes a site for future polymer initiation, so the array can maintain a steady-state polymer number. In contrast, the catastrophic disappearance of a microtubule that is free in solution removes a pair of ends from the reaction mixture, so the number of polymers continually decreases as the length of those remaining increases.

The model laid out by Mitchison and Kirschner has interesting implications for the behaviour of microtubules *in vivo*, particularly for the fibres of the mitotic spindle. Their work suggests that the lability of the spindle is due to a rapid disassembly of some microtubules, the slow growth of others, and a continual re-initiation of new microtubules from the centrosomes. The length distribution of the microtubules would then be based on the actual statistics of the GTP-tubulin caps.

While it has been known since the work of Inoue in the 1950s that spindles are labile *in vivo*, the extent of spindle dynamism has recently been shown to surpass expectations¹² and even to defy explanation by conventional wisdom about microtubule dynamics. Papers in the December issue of the *Journal of Cell Biology* by E. Salmon *et al.* and W. Saxton *et al.* will describe both the rates of incorporation of fluorescent tubulin microinjected into living mitotic cells, and the rates of fluorescence redistribution after photobleaching of fluorescent tubulin that is already equilibrated with cellular pools. The half-time for spindle microtubule turnover seems to be 10–20

seconds, a rate too fast to be explained by an end-exchange of subunits as conceived by Oosawa and Kasai. While the real explanation of these fast rates is not yet known, the work by Mitchison and Kirschner offers a testable model.

Their model for microtubule behaviour fits with one current concept of spindle formation — that some of the microtubules initiated at the poles are captured by kinetochores and thereby selectively stabilized. It also fits the observation that spindle microtubules, which interact with one another at the spindle midplane where fibres from the two poles interdigitate, are stabilized and can even elongate, while shorter microtubules, which fail to reach the zone of interdigitation, disassemble¹³.

Mitchison and Kirschner seem to have hit on something fundamental about microtubule behaviour. It will be interesting to follow the direct test of their suggestion that the hydrolysis of the GTP-tubulin cap is responsible for the behaviour they see. Because the polymerization rate

constants of actin show a similar dependence on the hydrolysis of bound nucleotide¹⁴, it will be interesting for students of actin microfilaments to ask how much of all this thinking is also relevant to the *in vivo* behaviour of their polymer. □

1. Mitchison, T. & Kirschner, M. *Nature* **312**, 232 (1984).
2. Mitchison, T. & Kirschner, M. *Nature* **312**, 237 (1984).
3. Brinkley, B. *et al.* *J. Cell Biol.* **90**, 54 (1981).
4. Roopal, A. *et al.* *J. Cell Sci.* **55**, 365 (1982).
5. Kuriyama, R. *J. Cell Sci.* **66**, 277 (1984).
6. Oosawa, F. & Kasai, M. *J. mol. Biol.* **4**, 10 (1962).
7. Margolis, R. & Wilson L. *Cell* **13**, 1 (1978).
8. Carlier, M.-F. & Pantaloni, D. *Biochemistry* **20**, 1918 (1981).
9. Hill, T.L. & Carlier M.-F. *Proc. natn. Acad. Sci. U.S.A.* **80**, 7234 (1983).
10. Carlier, M.-F., Hill, T.L. & Chen, Y.-D. *Proc. natn. Acad. Sci. U.S.A.* **81**, 771 (1984).
11. Hill, T. & Chen, Y.-D. *Proc. natn. Acad. Sci. U.S.A.* **81**, 5772 (1984).
12. Salmon, E. *et al.* *J. Cell Biol.* **99**, 1067 (1984).
13. McDonald, K. *et al.* *J. Cell. Biol.* **83**, 443 (1979).
14. Pollard, T. *J. Cell Biol.* **99**, 769 (1984).

J. Richard McIntosh is in the Department of Molecular, Cellular and Developmental Biology, University of Colorado at Boulder, Boulder, Colorado 80309, USA.

Conservation

Pandering to Western pandas

from Brian Bertram

THE captive giant panda population in the West is in dire straits, consisting of only 14 animals. Apart from one baby, a mere four of them are females; they are past middle age, and only two of them are reproducing. On the other hand, there are four good unpaired males. Nonetheless there has been a dramatic improvement in the captive breeding of giant pandas in the West recently, with seven babies born over the last five years. A symposium at West Berlin Zoo on 29–30 September 1984 brought together scientists and managers from all European and North American countries



that hold the species, to share their experience and co-ordinate future action.

It became clear that the large amount of research into giant panda reproduction, perforce in captivity, has been highly productive. Recognition of peak oestrus, by both behavioural and hormonal signs, is becoming more accurate. Pregnancy can now be more reliably assessed, although not until after what seems to be about three months delay in implantation, and birth dates better predicted. The very strong maternal behaviour is becoming better

understood, so techniques for artificial assistance have developed rapidly. Semen collection is almost routine and long-term semen storage is becoming more practicable. Artificial insemination techniques have benefitted from modern equipment and from recent experience of the detailed anatomy of the female giant panda.

In captivity, the species appears to be particularly vulnerable to intestinal and renal disorders. Dramatic major illnesses have developed rapidly in different collections; about half the cases have been remedied but only through rapid expert and collaborative attention.

The extent of co-operation between the keepers of captive giant pandas has been unprecedented. Within the past four years, there have been six international transfers, involving semen for artificial insemination, blood for emergency transfusions, and a live animal for mating. Veterinary staff and advice have flowed freely between institutions. Semen banks and tissue stores are being established and cell lines set up. The international studbook for the giant panda, one of many organized under the auspices of the International Union of Directors of Zoological Gardens, collects and collates information on giant panda matters. Considering the critical position of wild giant pandas, the establishment of a viable captive population is imperative. □

Brian Bertram is international studbook keeper for the giant panda and Curator of Mammals at The Zoological Society of London, London NW1 4RY, UK.

Palaeomolecular biology

Raising the dead and buried

from Alec J. Jeffreys

Is the quagga as dead as a dodo? Not entirely, and nor indeed might be the dodo, if the remarkable findings of Russell Higuchi, Allan Wilson and co-workers reported on page 282 of this issue are anything to go by. Because even though the quagga, a curious chimaera of horse and zebra, became extinct just over a century ago, some of its DNA has survived in a museum specimen in a state suitable for molecular cloning.

As a start, Higuchi *et al.* managed to isolate partially degraded DNA from dried quagga tissue and to show that at least some of it is of quagga origin, rather than a contaminant, by virtue of its hybridization with DNA from the closely related zebra. More importantly, it was possible to obtain clones of specific mitochondrial DNA sequences. Comparison of these sequences with those of zebra mitochondrial DNA provided the final proof that they are of quagga origin. Not only are the two species' sequences as closely matched as is expected for congeneric animals but most or all of the differences that do exist (mainly synonymous base transitions at third codon positions) are clearly not due to postmortem changes in DNA that have been subsequently misrepaired during cloning, and so they must reflect the evolutionary history of the quagga. Clearly, the great power of molecular phylogenetic analysis, so far restricted to living animals, can now be brought to bear on at least some extinct species.

The choice of mitochondrial DNA was wise in view of its abundance: there are many mitochondria in each cell. However, to extrapolate from Higuchi *et al.*, it might even be possible to extend such studies to single copy genes of the nucleus. Thus 3 gm of preserved tissue should yield sufficient DNA to produce a library of up to 3×10^7 clones, each containing perhaps 100 base pairs of quagga DNA and together covering most of the genome.

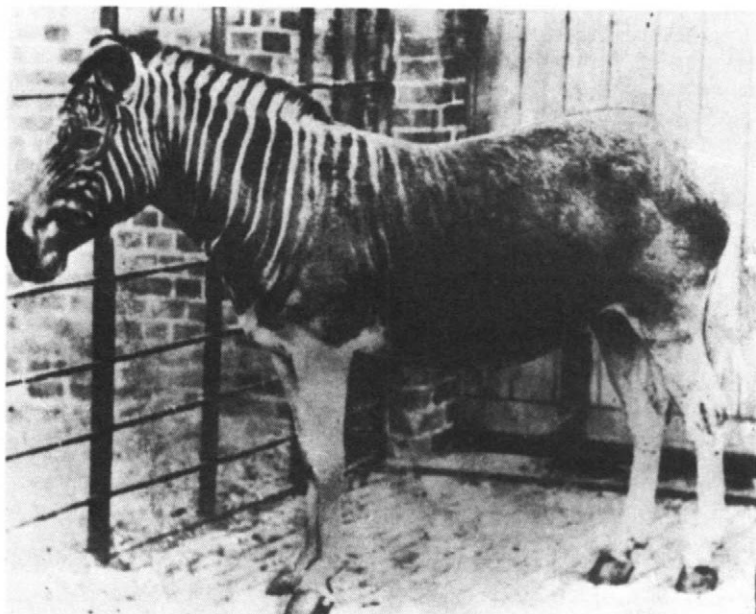
The obvious next question is whether other museum specimens will yield up their molecular secrets. One can only hope that museum curators will be reasonably sympathetic to hordes of molecular biologists eager to dismantle their cherished exhibits. Anthropologists could benefit too — DNA sequences of bog people and Egyptian mummies would no doubt be fascinating, though cloning the latter might prove too nerve-racking a task for the superstitious genetic engineer.

A century or two in a museum is one thing; 40,000 years in a Siberian permafrost bog — the fate of the Magadan mammoth — is another. Nevertheless, Wilson's group (*Fed. Proc.* 43, 1557; 1984), M. Goodman's (*Acta Zool. Fennica*, in the

press) and my own (unpublished) have shown that substantial quantities of DNA can be recovered from preserved mammoth tissue. Unfortunately, almost all of it comes from recent microbial contamination, probably introduced after excavation. Elephant-like DNA sequences are present at vanishingly low levels (less

evolutionary synthesis by studying fossil DNA, still look like nothing more than a glorious dream. However, it is far too early to give up, and it might just be possible that DNA has survived in some fossilized material.

One final point: DNA can easily be purified from animals that have died, and once dried it is stable and should survive for centuries without degradation. It is therefore vital that zoos or museums should start systematically to store DNA from as many species as possible, and certainly from any that face extinction. Friedrich Miescher, who discovered nucleic acid in



A quagga mare exhibited at London Zoo from 1851 to 1872. The photograph was taken 13 years before the death of the last captive quagga in Amsterdam in 1883. (Zoological Society of London.)

than 1 part in 10^4 of the total DNA) and are severely degraded. Cloning this DNA would indeed be a mammoth task, and any sequence information recovered would probably be seriously distorted by post-mortem modification. We know nothing about the chemistry of DNA degradation over geological time periods.

Any hopes that molecular biology and palaeontology can be fused into a grand

1868 (see Vogel, F.C.W. *Die histochemischen und physiologischen Arbeiten von F. Miescher*, Leipzig, 1897), could have saved Higuchi *et al.* a lot of trouble if he had had the foresight to make and store fresh quagga 'nuclein'. □

Alec J. Jeffreys is in the Department of Genetics, University of Leicester, Leicester LE1 7RH, U.K.

Particle accelerators

New concepts for high energies

from J.D. Lawson

SPECTACULAR development of particle accelerators has sustained the progress of elementary particle physics during the last half century. Although current projects, and others yet to be funded, will ensure the progress of particle physics until the end of the century, there is serious concern about what will happen beyond then. We need some new concepts or some radically new technology to continue past trends. Two years ago high energy and accelerator physicists met in Oxford to consider 'The Challenge of Ultra-High Energies'. A few

weeks ago they met again, this time in Frascati*, to examine in more depth some of the ideas discussed at Oxford.

The 'colliding beam' concept, first demonstrated twenty years ago, becomes particularly effective when γ , the ratio of total particle energy to rest energy, becomes large. Almost the total energy of $2\gamma m_0 c^2$ is then available for particle creation and excitation, whereas if one of the particles is initially stationary in a target, it

*The Generation of High Fields for Particle Acceleration to Very High Energies, Frascati, 25 September - 1 October 1984.

carries away most of the energy as kinetic energy and only $\sqrt{2}m_0c^2$ is accessible. The large values of γ now available, it seems that future high energy accelerators must operate in the colliding beam mode. The large electron-positron (LEP) storage rings now under construction at CERN near Geneva represent the most advanced enterprise of this type. The tunnel to house the rings is 27 km in length, and the planned energy is ultimately expected to exceed 100 GeV per beam ($\gamma = 2 \times 10^5$). An even larger machine, for protons of energy 10-20 TeV per beam, is being considered in the United States. These machines are near the limits of acceptable cost and sheer physical size. In the case of electrons, the energy loss by synchrotron radiation in the magnet fields also represents a technical limit; further progress will require 'linear colliders', in which two linear accelerators fire very dense bunches of electrons at one another. An experiment to test the principle of this method, and to learn how to generate and control the very dense bunches that will be needed, is underway at the Stanford Linear Accelerator Center (SLAC).

Two questions were asked at Frascati: first, what is the minimum performance specification for an 'interesting' accelerator beyond the range of established techniques; second, what approaches might ultimately be capable of achieving the necessary specification? Appropriately, Carlo Rubbia (CERN) opened the discussion by attempting to define not only what energies are required, but what minimum rate of worthwhile events would need to be achieved.

The ratio of event rate to cross section is known as the luminosity, which for pairs of bunches each containing N particles colliding head-on is given by the expression $L = N^2 f / A$, where f is the repetition rate and A the bunch area. For linear colliders the bunches are used once only, so that the beam power is $P = N f m_0 c^2$. L is determined by the minimum acceptable event rate for experimentation, and P by the power supplied to the accelerator times the efficiency of acceleration of the beam. Since cross sections for interesting events involving point-like particles decrease with energy, both L and P must increase, and Rubbia made a plea for $L = 10^{34} \text{ cm}^{-2} \text{ sec}^{-1}$ for a 5 + 5 TeV electron machine. Inserting affordable values for P and technically-credible values for f and N in the above formulae yields sub-micron diameter bunches with an optical beam quality that is better than has so far been achieved. This challenge was quantified by L. Hand (Cornell University) and J. Rees (SLAC), who indicated the need for high efficiency and optical precision in the beams. High fields are also essential, and it was this aspect that was emphasized at the meeting. Recent experiments at SLAC suggest that extrapolation to above 100 MeV/metre should be possible by development of conventional techniques; this implies not more than 50 kM for a 5 TeV machine, a large

but not incredible value. Any new approach must either enable such fields to be obtained more economically than at present or promise higher fields.

Progress was reported on two other linear accelerator concepts, the 'wake field' accelerator of G. Voss and T. Weiland (DESY) and the 'two beam' concept of A. Sessler (Lawrence Berkeley Laboratory). In the former, converging wake fields from a ring of electrons moving through a suitable structure produce intense fields on the axis which can then be used for acceleration. Experiments to test the idea are under way at DESY in Hamburg, FRG and, with a different geometry, at Osaka in Japan. In the two beam accelerator power is provided not by an array of klystrons, but from a single low energy electron beam that runs parallel to the accelerator structure. This passes alternately through magnetic undulators, which generate the required microwave power by free electron laser action, and induction sections that restore the beam energy. An experiment at Livermore has produced 100 MW of radiation at 34 GHz with 5 per cent efficiency. Both these ideas aim at high fields, but their potential still needs to be established.

As at the previous meeting, the 'beat-wave' concept of T. Tajima (University of Texas) and J. Dawson (University of California, Los Angeles), excited considerable interest. It promises the most spectacular accelerating fields, generated by beating together two laser beams in a plasma, where the difference of the laser frequencies is equal to the plasma frequency. Despite tantalizing figures for the field gradient, the detailed physics of the process is still not clear and it is too early to see how one might design such an accelerator. In the United States, experiments are planned to try setting up

and detecting a beat wave, and there are hopes that a experiment will be approved in the United Kingdom. There is plenty of scope for further theory and simulations, and for ideas on how to use the concept to build an accelerator. The stringent requirements for efficiency and good beam quality make this task difficult, but there was plenty of enthusiasm for further exploratory work.

An idea much discussed in earlier meetings is the acceleration of particles over a grating carrying an evanescent surface wave, excited by an incident laser beam. At its simplest this would not be efficient, but basic experimental studies on the coupling of waves and grating structures are being planned in Ottawa. A development due to R. Palmer (Brookhaven) is to replace the grating by two rows of tiny spheres spaced about three to a wavelength. The idea is to squirt these from an ink jet printer and, when they have arrived at the correct position, to subject them sequentially to a travelling pulse of laser light. This should render the spheres conducting and, on the short time scale required, capable of sustaining the enormous fields required for rapid acceleration to high energies of a beam of injected particles. The spheres act as a sort of receiving antenna that transfers energy from the laser light to the particles.

Everyone will have their own views about how credible these new concepts really are, but there was considerable enthusiasm at the meeting and no doubt that the future is going to bring increasing activity in this fascinating and potentially important field. □

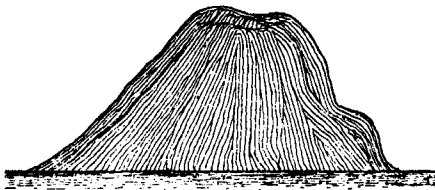
J. D. Lawson is at the Science and Engineering Research Council Rutherford Appleton Laboratory, Chilton, Didcot, Oxfordshire OX11 0QX, UK.

100 years ago

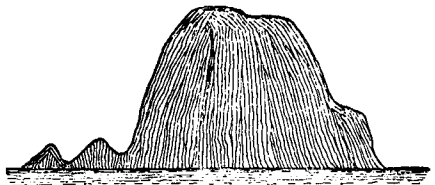
THE NEW VOLCANIC ISLAND OFF ICELAND

At the end of July this year the light-keeper at Cape Reykjanes, the south-west point of Iceland, reported that a volcanic island had risen in the sea a few miles off the cape. Reykjanes has long been noted as a centre of volcanic activity, and from time to time islands have arisen and submarine eruptions have occurred in its neighbourhood. In the year of the great Skaptárfell eruption, which proved so fatal to Iceland, 1783, an island appeared off Reykjanes, only to disappear again after a very brief existence. Only a year or two ago an eruption of considerable violence occurred in the sea, not far from the spot where the new island appeared. Columns of steam and clouds of dust, mingled with occasional glowing masses of fused rock, were seen to rise out of the sea, and large quantities of pumice were thrown up and drifted ashore on the neighbouring coast.

When first seen, on July 29, its shape was that of a truncated cone with a slight depression on the top, and a considerable hollow half way down the slope on the north side. On August 5 and 6 a series of violent earthquake shocks occurred, which shook and split the masonry of



The island as when first seen.



The island as it now appears.

the lighthouse and damaged the lamps. The island's shape was considerably altered: a large part of the slope on the south side had slipped down into the sea, where it now lies, forming two little hillocks close to the foot of the main mass, and leaving a steep face nearly perpendicular towards the bottom.

From *Nature* 31, 37, 13 November 1884.

Igneous petrology

Old magma chambers replenished by new ideas

from R. Stephen J. Sparks and Herbert E. Huppert

IGNEOUS processes are a dominant mechanism for chemical differentiation within the Earth. They set up complicated plumbing systems which connect the deep source of magma generation to the surface. Some magma ascends to the surface to initiate volcanic activity, while some can be trapped at depth to form intrusions. The plumbing can include magma chambers, which for some time have been envisaged to play an important role in controlling the chemical evolution of magmas and the behaviour of volcanoes. Only a decade ago, magma chambers and their associated conduits were generally assumed to be closed systems in which chemical equilibrium was commonly attained. A recent meeting* illustrated how ideas have changed; the main theme was that most magmatic systems are open, being replenished from below and tapped from above with non-equilibrium and dynamical processes playing an important and dominant role.

New ideas on the role of fluid dynamics in the evolution of magma chambers formed a major part of several presentations. These included simulations carried out in our laboratory of the convective movement of magma through a crystal pile, which has been proposed as a mechanism for adcumulus growth (*Lithos* 17, 139; 1984). The fluid dynamical principles of magma chamber replenishment by both heavy and light inputs of primitive magma were described by a number of speakers who argued that thermal and compositional stratification can be created by such replenishments, as well as by mixing between magmas. Illustrating the practical application of such ideas, A. Naldrett (University of Toronto) explained how a light input of primitive magma can lead to the formation of PtS ore deposits in the Bushveld Complex (*J. Petrol.* 24, 133; 1983). S. Blake (Australian National University) discussed the eruptive withdrawal of a two-layer chamber and suggested that only very large eruptions can tap magma from a lower layer which is more than a few tens of metres below the chamber vent.

One of the most fascinating and controversial topics in recent petrological discussions concerns the origin of compositional zoning in magma chambers. Some detailed petrological studies (G. Worner, Bochum University and J. Wolff, University of Texas) showed how extreme fractionation of trace elements can occur in

alkaline phonolitic systems as well as in the better-known rhyolitic systems. Some elegant experimental studies on Soret diffusion (C. Lesher, Yale) showed that distribution coefficients of trace elements are very sensitive to slight changes in melt composition in highly evolved magmas. This implies that large changes in melt structure accompany fractional crystallization of highly evolved magmas and strongly influence the development of extreme enrichments or depletions of trace elements. A. McBirney (University of Oregon) described the results of fluid dynamical analyses of boundary-layer flows in a magma chamber which is cooling and crystallizing at a vertical wall in an otherwise uniform magma. He showed that many magma chambers should be in the regime of counter flow, in which a thin, inner compositional boundary layer of differentiated fluid can ascend to the top of the chamber to form a zoned region. The theory predicts a vertical flux of $1 \text{ cm}^2 \text{ s}^{-1}$ per unit wall width, somewhat larger than previously postulated. The mechanism provides the most plausible way yet proposed to cause zonation in highly evolved magma chambers.

Volcanic studies

New ideas in experimental petrology are contributing to our understanding of open magmatic systems. T. Grove (Massachusetts Institute of Technology) showed how small amounts of crystal contamination can have substantial effects on the chemical evolution of magmas, producing calcalkaline rather than tholeiitic trends from a basaltic parent. An application of experimental and observational petrology to the Mount St Helens magma from the 18 May, 1980 plinian eruption established that the magma had a temperature of about 920°C , ascended from a chamber at 6–7 km depth and contained 4.6 per cent H_2O (H. Sigurdsson, Rhode Island and M. Rutherford, Brown). The petrological model showed excellent agreement with volcanological and geophysical results.

A major reason for believing open-system behaviour is dominant in many magmatic systems has come from detailed investigations of the petrology and geochemistry of igneous centres that have well-documented geological histories. The fine details of individual volcanic and plutonic units have shown that closed-system processes generally provide inadequate explanations of the data. Several speakers discussed how an interaction of basaltic magma with crustal rocks is required to

explain the isotopic systematics. For example, volcanic rocks of the Hebridean Province, Scotland (R. Thompson, Imperial College) and the Taos Plateau, New Mexico (M. Dungan, Southern Methodist) show isotopic and trace element signatures of complex interaction with the crust. Z. Palacz (Open University) documented how large variations of isotopic composition can occur within individual layers of the Rhum intrusion, which implies that open-system processes play a dominant role.

An interesting controversy emerged from petrological studies of the Fish Canyon Tuff (J. Storer, Rice University). The conventional view of the evolution of large caldera complexes envisages establishment of a high-level magma reservoir prior to collapse. Compositional zonation is thought to develop in such chambers. Storer, however, concludes that the Fish Canyon magma is not chemically zoned and that the minerals in the magma formed in a deep chamber, 20–30 km in the crust. If a shallow chamber existed before the eruption, then it must have been intruded rapidly, since there is no signature of its existence in the mineral chemistry. This is still very controversial but, if correct, the arguments for the conventional high-level magma chamber will have to be re-examined.

Isotope geochemistry, which has led the way in recent years on many questions of magma genesis, is now producing a deluge of data, combining a number of radiogenic and stable isotope systems. For example, several speakers showed how the sediment, mantle and crustal components can be distinguished in island arc volcanics. However, there are signs that the law of diminishing returns may be operating. The answers to some fundamental questions (for example, how much subducted oceanic crust and sediment is recycled in arc magmatism) remain essentially qualitative. More and more isotopic data may not produce any clearer picture unless combined with other approaches.

Finally, everyone was greatly saddened by the death of the Japanese petrologist Masanori Sakuyama, who was to have presented a paper at the conference. Sakuyama was an outstanding young scientist who had already made several important contributions to understanding open-system magmatism, particularly the role played by magma mixing and andesite petrogenesis. □

R. Stephen J. Sparks and Herbert E. Huppert are supported by the BP Venture Research Unit in the University of Cambridge's Department of Earth Sciences, Cambridge CB2 3EQ, and Department of Applied Mathematics, Cambridge CB3 9EW, UK.

Erratum

In Fig. 2 of Nancy L. Craig's article 'Resolution of intermediates' (25 October, p. 707), the labels *attL* and *attR* were misplaced and should be transposed.

*'Open Magmatic Systems', Fort Burgwin Research Centre, New Mexico, 21–28 August, 1984, sponsored by the Institute for the Study of Earth and Man at the Southern Methodist University.

Meeting in print

Conference proceedings — symposia — are generally disliked, yet make up a large part of the lists of most scientific publishers. Anthony Watkinson examines this aspect of publishing.

PUBLISHERS have their curious in-house technical terms, perhaps to bemuse authors, and when they write about symposia they mean conference proceedings and not, as was once the case, any variety of multi-authored book. To put it mildly this type of publication has a bad name. Symposia are a major vehicle of scientific communication, but with one or two notable exceptions, such as the series emanating from Cold Spring Harbor Laboratories, they are generally condemned and disliked. So why do publishers find them so attractive?

Commissioning editors are familiar with the fact that when they seek advice from an academic about publishing a symposium the adviser often seems to be writing in reply as if quoting from a partially remembered script. Unreserved condemnation is frequent:

There are far too many symposia published. Like the others, we are likely to get from this one a book which will be out of date before it is published and which will be both uneven in level and patchy in quality. It should be rejected.

Where there is approval it is grudging:

Although I am against such publications as a matter of principle, I am prepared to make an exception in this case where there is a real need for a state-of-the-art volume in this very important (to be interpreted — *my own*) field. It is rapidly growing (*my research group managed to get out five papers last year*) and is attracting a lot of interest. I must also say that the organizer has managed to get together the leading people (*including me but that does not influence my judgement*) and on balance I am inclined to recommend that you take this one on.

These are generalizations, seen through the distorting glass of publishing preoccupations. The designation "symposia" actually covers a wide range of types of books and some demarcation within the genre is necessary. A convenient criterion is the nature of the (type) setting. Does the book represent the summit of the printer's art — what used to be letterpress and art paper — or crummy old author-prepared camera-ready copy?

The heavy end of the range is best represented by those old-established series produced for British learned societies, characteristically in classical biology. A very high standard of production is expected. Line figures are even redrawn and relettered to achieve conformity — the ultimate in luxury now that the drafting staffs of publishers have been cut to the bone — and much care is lavished on micrographs of sections of voles' testicles and suchlike. Usually pressure to get the book out

quickly is not overwhelming, and in general it is as a collection of authoritative reviews that such symposia are judged.

Whereas such series are often cited as serials (symposia of this society and proceedings of that), there are other typeset review volumes which are presented as if

Even where the book is admitted to be a symposium the words "based on" or "derived from" coyly suggest something else.

the origin of the book were a shameful secret to be relegated to the preface. Even where the book is admitted to be a symposium the useful words "based on" or "derived from" coyly suggest something else; and it is quite common to choose a more challenging and supposedly "selling" title (*The Worm in the Bud*) when the subtitle reveals that we are concerned with the third international meeting of the society for weevil watchers.

It is not uncommon for the multi-volumed proceedings of the big disciplinary meetings (for example the international get-togethers of immunologists) to be "properly" typeset, after having been seriously edited, while at the same time coming out within six months of the actual meeting. In older disciplines both the product of such proceedings and the function of the meeting of the international union itself are less highly regarded. To manage the schedule prodigious organizational capacity is required both of the publications committee of the learned body and the publisher; a team of desk-editors, who usually have their being in windowless

rooms, may be plucked from their desks and flown to Kyoto by an apparently generous employer, there to spend their days and nights copy-editing.

Another category of typeset symposia is the "workshop" publication favoured by the drug houses and financed by them. These are prestige jobs and they must look like "real" books. The size of the print run and the amount of money on offer can stimulate even the most sluggish of book publishers to manage the sort of schedules which are normally possible only for journals.

Most symposia don't look like "real" books but are produced from camera-ready copy. Until recently this meant typewriter-produced script with ragged (unjustified) right-hand margins. Quite a few publishers will only "do" meetings if the editor/organizer agrees to this route. It is one of those house rules which indicate that the management has got so far away from line responsibilities that they have to rely on arbitrary criteria. It exists because those in charge want to avoid investment in composition, not because they have any wish to insist on old-style typescript as such. If the volume is to be "re-typed" after editing under the auspices of the editor/organizer, or through the institute or society concerned, the product can now look like, and indeed is, cheap printing *sensu stricto*. Laser printers have become commonplace in the publication units of the wealthier institutes, such as those that house the "big" physicists, and even mere universities aspire to them. Even where the machine being operated is still actually called a typewriter, the typescript can be right-hand justified and elaborately set up, by the manipulation of golf balls, in italic or bold as appropriate (or sometimes not). It does not matter that researchers in applied psychology have convincingly shown that unjustified typescript is easier to read than its badly justified equivalent. Publishers know what looks right. ▶



Talk among the posters. What will result from the talk inside the conference hall?

Finally at the bottom end, as it were, are those numerous symposia produced from author-generated camera-ready copy. These are the norm, the least regarded, the most dismissed: more on them later.

Publishers don't boast about the symposia on their list. A standard publishing house in science, technology and medicine (STM) will have at least 25% of its annual output of new titles in this category, but such books will be the least promoted. In some disciplines, however, for example the neurosciences, a new publisher can build up a front-running list on little more than well-chosen proceedings and for the new branch of an existing international firm it could be argued that symposia publishing is the quickest and easiest way of getting in on the act. Many journal publishers claim to publish symposia to keep their journal editors happy. Other publishers could as reasonably claim, but don't, that the efficient publication of a meeting for a key person can put them in the running when a journal is hatched.

Publishers don't always themselves recognize that they make money out of symposia, but they do. The economics of the quick and dirty end of the spectrum

Publishers don't always recognize that they make money out of symposia, but they do.

illustrate this fact most clearly. In the first place, if there are no composition costs investment is minimized. It is hardly a major outlay to supply the familiar blue-lined paper (the grid) and the subtly differing IMPORTANT INSTRUCTIONS FOR YOUR TYPIST (the rest of the package). Secondly it is often the case that this material is sent out by the academic editor himself or herself and it may also be the editor's job to make sure that the contributors sign the minimalist contract to establish that the publisher really does hold the rights. Whatever the hassles that will arise, they are for the editor to sort out, not the publisher.

The third advantage is the convention over royalties. Whereas 15% of net receipts or thereabouts is regarded as reasonable in the case of a monograph, and the sum of royalties and fees for those involved in a specially commissioned multi-authored book is much of a muchness, editors of symposia cannot hope to be propositioned with inducements like this. Ten per cent of income is tops in most cases, going to the editor/organizer only, and generally it is considered perfectly appropriate (given a buyer's market) to screw down symposium royalties to purely token levels where for other categories of book such an action would be regarded as at best unsuitable and possibly even counter-productive. All this is good for the business arithmetic. It must however be said in qualification that where major international and prestigious national learned societies are concerned, contracts involving royalty arrangements running to as much as 13% of the list price

are known to exist and experienced treasurers (biochemists are good at this) have secured rather large down-payments.

Fourthly there is a further financial consideration which endears symposia to the hearts of publishing accountants. It is one area in which their fantasies can have free rein because it is generally agreed that symposia are not price-sensitive items. In theory publishers price to a market; textbooks, for example, cannot be higher than a price determined by that of the competition. In practice, although types of book mainly bought by libraries may be priced in accordance with arcane formulae that do not relate to costs — bringing into play such concepts as price ceilings (if the book is over \$100, will it go to the library committee?) — publishing editors fight the rules to keep down the price of their own creation and someone else's life's work as a scholar. No one cares about symposia, especially the camera-ready variety. From some houses amazingly high prices — when costs are considered — and absurdly low breakevens are par for the course. Production costs may be covered if only 100 are purchased. An examination of many STM catalogues to compare the price per page of a typeset monograph and a symposium volume in the same field that has been produced from author-generated camera-ready copy may show little difference between the two categories.

Other advantageous arrangements are possible. Publishers may lash out with cash on advances and composition where the big meetings are concerned, but though they may talk of generous donations there is method in their madness. The publication is a very visible one. Spin-off books and even journals may result. And in the short term there is a good chance that the price of the book-of-the-meeting may be built into the registration fee so that costs are covered before publication. Even where the publisher is only permitted to make a special offer to participants for these sort of meetings, usually in a newish and still-cohesive field, experience shows that most of those attending do actually order the book or books at the special price and those that could not actually make it to the occasion feel that they should have the book on their laboratory shelf.

Publication for the pharmaceutical industry is a specialist area, shrouded in mystery and not to be touched on lightly; exposure of the secrets is probably punishable under Swiss Law. The subsidizing of such publications, though significant to the publishing house, represents a very small part of the marketing budget of a big drug company. The marketing manager of such an organization has simple needs; he wants a book that comes out on time, under a good imprint, and looks like a book. The publisher quotes a fee which is estimated on costs per page (plus, plus, plus) and, if his print buying is good, profitability is ensured before a copy is sold.

Actually, the most prestigious category

of symposium volume is just about the least profitable category of book that the scientific publisher can take on (except for the dreaded *Festschrift*). The books are expensive to produce but the officers of the learned society prevent "realistic" pricing. Profits are difficult to achieve and payback is slow. If the topic of the meeting looks

...if the organism concerned is small and has lots of legs the bottom line is pretty sure to look bad.

boring to the layperson it probably is for most scientists also. It would be invidious to pick out particular topics which chill the hearts of publishers, but if the organism concerned is small and has lots of legs the bottom line is pretty sure to look bad. The author who wonders how the publisher can afford to embrace the series of the —ology Association when his own important opus has been rejected as uncommercial should turn to the journals section of the same catalogue. It is a matter of sprats and mackerels.

A publisher writing in *Nature* must take it that publishers are in the symposium business for the cash, the contacts or both. What of the scientists' point of view? The transformation of a transient three or four days of academic exchange into an enduring collection of reviews, or a published state-of-the-art forum, can be an invaluable contribution to learning. And where a subject is emerging as a discrete area of research activity, where results are coming thick and fast or theories are presented and overturned month by month, meetings are crucial to the scientific community and their record in printed form is equally essential. In such circumstances, where something more discursive and less restricted than the scientific paper is required, and some sort of synthesis, albeit tentative, must be attempted, symposium proceedings are the most useful and probably the only practicable form of written communication.

There are other occasions where symposia are to welcomed but not always. In the past there was always some publisher at hand to take on any meeting however unappetizing it might be. It is said now that some proffered volumes are rejected by all. One might hope too that more organizers will look to the Harden and Gordon model and not seek volume publication as a reflex action, and that more contributors who have given their all several times already will refuse to submit a paper when another unexpected demand for the same material is made. And if publication is accepted to be a good idea by everyone concerned, perhaps publishers may consider treating symposia less as the cannon fodder of the book industry but rather as books with a purpose to be cherished and made the best of. □

Anthony Watkinson is a Commissioning Editor at Oxford University Press.

The light and the dark

Lewis Thomas

The Limits of Science. By P.B. Medawar.
Harper & Row: 1984. Pp. 102. \$11.50.

To be published in Britain in February 1985 by Oxford University Press.

THESE are hard days for science, for all the achievements of the past two or three centuries that have made possible the lives of vast numbers of human beings who would never, without the works of science, have survived the population curve that now seems to be turning into a sustained near-vertical upswing. Food production on a scale sufficient to have kept most such populations from crashing long since, medical and public health advances that have nearly doubled the possible life expectancy in affluent nations in just this century, technologies for communication that ought to be offering the prospect of world-wide education for all people, not to mention diversion and entertainment — these are among the contributions of basic scientific research and its companion art, applied science. More than these, science's greatest contribution to humanity has been the beginning of something like a comprehension of the true nature of Nature, in a range extending from the most intimate and essential transactions within a single living cell down to the fundamental nature of atoms and their particles, and out to the furthermost reaches of the Cosmos and, perhaps, some of the detailed events of the birth of the Universe itself.

No doubt about it: science is, as Peter Medawar asserts in the preface to his latest and, in my view, most wonderful book, "a great and glorious enterprise — the most successful . . . that human beings have ever engaged in".

But it is not seen this way by everyone, most certainly not by a great many thoughtful, well-educated, worried men and women in the streets of the world. To the contrary, science is widely perceived as the root cause of our most intractable problems: the population explosion itself, the detachment of humankind from all the rest of nature, the despoliation of the environment, the loss of any sense of mystery about the world and the way it works, the flattening out of human experience. And now, since the Second World War, the worst of things ever done by science, the overhanging threat to the planet of thermonuclear weaponry.

The *Limits of Science* does not settle the dispute, but it does do something almost as important, perhaps more important than winning an argument. It lays out the issues at stake, fairly and gracefully. Medawar is wholly on the side of science, he gives no

ground, he is determined and firm in his views. In the first third of his book, a chapter titled "The Nature of Scians", he moves freely from one large topic to another, briefly and aphoristically, the whole intended to provide the reader with information needed for an understanding of what science really is, how it is carried out, and what makes it work when it luckily does work. In one paragraph headed



"Science and Politics", he writes,

Some of my colleagues will think me sensible, others disloyal, when I roundly declare that political and administrative problems are not in general scientific in character, so that a scientific education or a successful research career do not equip one to solve them More than that, I regard my opinions as so obviously right as not to be worth the exertion of justifying them.

But then, he goes on justifying in several pages of historical instances in which politicians and statesmen should have found scientific knowledge indispensable for making decisions affecting the public welfare, including this one:

I have never met a scientist who does not believe that the effects of ionizing radiations from a nuclear bomb, especially upon deoxyribonucleic acid, the vector of genetic information, are so utterly frightful as to outweigh all other considerations in framing national policy, among which I include national pride (the chap inside is past caring whose flag flies over his tomb).

I must say that I have run into a scientist or

two or three lacking such belief, but I agree that they are the rare exceptions.

The first two of the three chapters that comprise this compact book are vintage Medawar, so much fun to read that you may fail to recognize, until later, that he is writing seriously on serious matters, doing his best to protect scientific endeavour against damage by people (and governments) who have an insufficient grasp of the importance of the enterprise. He keeps coming back to his main point: "In terms of the fulfillment of declared intentions, science is incomparably the most successful enterprise human beings have ever engaged upon".

And then, in the last 40 pages, he moves into new, darker territory. What are the questions that lie forever beyond the reach of science, for which the answers can never

be stated flatly in scientific, testable, verifiable or falsifiable sentences? They are, he says, the questions about first things and last things: How did everything begin? What are we all here for? What is the point of living? Is there a God? Queries such as these, says Medawar, are no part of science's business. Not that they cannot be answered, he concludes, just that the answers are by their nature matters of faith, not matters accessible to reason.

He loses me, or I him, in his last pages. He defines "religion", and then sets it aside, as a set of explanations that

appeal to the agency of a personal God — and I should have insisted upon the notion of a personal God, for nothing could be more wishy-washy or generally unsatisfying than to liken God to some kind of diffuse benevolence that permeates the material world.

I do wish that Medawar, one of the most spectacular biologists of the century, in possession of one of the clearest minds, more acutely aware than most of his colleagues of the immense amount of biological and physical science still out there, waiting to be discovered, filled with mystery because still so unknown, had not left it there. Stick around, I claim, keep an eye on surprise. Maybe not the *very* first things, or the *very* last things, but some very early and very late things are waiting to be tripped over, and tripped over most likely by science.

But Medawar succeeds anyway, despite my kind of doubt and misgiving. He takes up important matters, honestly and in high courage. His book should be read and welcomed by scientists, especially young scientists just starting out, and by a non-scientist public wanting a better understanding of how a scientific mind works in high gear. In short, by just about everyone I can think of. □

Lewis Thomas is President Emeritus of the Memorial Sloan-Kettering Cancer Center, New York.

For goodness sake

Christopher Longuet-Higgins

From Knowledge to Wisdom: A Revolution in the Aims and Methods of Science.

By Nicholas Maxwell.

Basil Blackwell: 1984. Pp.298. £19.50, \$29.95.

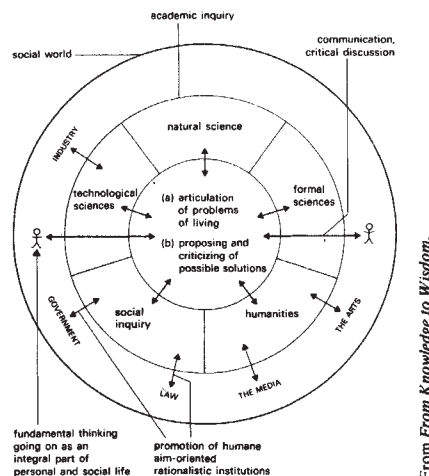
THIS book is the work of an unashamed idealist; but it is none the worse for that. The author is a philosopher of science who holds the plain man's view that philosophy should be a guide to life, not just a cure for intellectual headaches. He believes, and argues with passion and conviction, that the abysmal failure of science to free society from poverty, hunger and fear is due to a fatal flaw in the accepted aim of scientific endeavour — the acquisition and extension of knowledge.

It is impossible to do Maxwell's argument justice in a few sentences, but, essentially, it is this. At the present time the pursuit of science — indeed the whole of academic inquiry — is largely dominated by "the philosophy of knowledge". At the heart of this philosophy is the assumption that knowledge is to be pursued for its own sake. But the pursuit of objective truth must not be distorted by human wishes and desires, so scientific research becomes divorced from human needs, and a well-intentioned impartiality gives way to a deplorable indifference to the human condition. The only escape is to reformulate the goals of science within a "philosophy of wisdom", which puts human life first and gives "absolute priority to the intellectual tasks of articulating our problems of living, proposing and criticizing possible solutions, possible and actual human actions". The philosophy of wisdom commends itself, furthermore, not only to the heart but to the head: it gives science and scholarship a proper place in the human social order, whereas "standard empiricism" — the most articulate expression of the philosophy of knowledge — offers no kind of *justification* for the pursuit of knowledge, or, indeed, any basis for evaluating its results (that is, for distinguishing between a significant discovery and a trivial assortment of "scientific facts").

Nicholas Maxwell has breached the conventions of philosophical writing by using, with intent, such loaded words as "wisdom", "suffering" and "love". "That which is of value in existence, associated with human life, is inconceivably, unimaginably, richly diverse in character." What an un-academic proposition to flow from the pen of a lecturer in the philosophy of science; but what a condemnation of the academic outlook, that this should be so. Mr Maxwell is advocating nothing less than a

revolution (based on reason, not on religious or Marxist doctrine) in our intellectual goals and methods of inquiry. The foundation, in 1662, of the Royal Society of London for Improving Natural Knowledge was not, in that enlightened age, liable to serious misinterpretation; but in our own century the original purpose of the enterprise, defined by Francis Bacon as the enhancement of our power to do good, has been largely forgotten. "How elegant!" is a high compliment to a colleague's work. "How useful!" would be little short of an insult.

There are altogether too many symptoms of malaise in our science-based



Better relations? Interaction between the intellectual domain of inquiry and the social world according to the philosophy of wisdom.

society for Nicholas Maxwell's diagnosis to be ignored. He fair-mindedly admits that some scientists, of whom Albert Einstein was an outstanding example, do regard their work as a kind of pilgrimage — the goal of which is not mere knowledge but a deeper understanding of the natural world and of our place within it — and that such an attitude is characteristic of the philosophy of wisdom rather than the philosophy of knowledge. But it cannot be claimed that the best scientists have always been men and women of outstanding human qualities. Dr Strangelove is all too credible a villain, the more dangerous because idolatry — the worship of our own handiwork — is no longer regarded as a menace to society. (Would that the computers-in-education people took it more seriously than they seem to.) The dangers are less serious, though, in the more mature sciences, such as mathematics and physics, than in the much-newer social sciences — as their practitioners are pleased to call them. In Mr Maxwell's words:

The central, and tragic, intellectual mistake (according to the philosophy of wisdom) that has bedevilled social inquiry ever since the Enlightenment is illustrated in miniature in an especially graphic and simple way in a book by Barbara Wootton entitled *Testament for Social Science* (1950). Its subtitle — 'An Essay in the Application of Scientific Method to Human Problems' — might lead one to believe that the

book expounds and defends the philosophy of wisdom. But if the first few sentences of the book strengthen this belief, what follows must quickly dispel the idea. The book opens as follows:

"The contrast between man's amazing ability to manipulate his material environment and his pitiful incompetence in managing his own affairs is now as commonplace as it is tragic. . . I hope to show that the potential contribution of science in this field is far greater than anything we have seen: *the differences between the material of the social and the natural sciences [reviewer's italics] are differences of degree rather than of kind.*"

In the 34 years that have followed Barbara Wootton's *Testament* the social "sciences" have lamentably failed to make people any better because, suggests Mr Maxwell, they have not even tried; professional etiquette precludes a social "scientist" from making *value* judgments, from suggesting that anything might be *wrong*, from recommending any sort of *action* that might *interfere* with the system under observation. Mercifully these morally paralysing taboos have not yet overtaken the medical profession, who are commendably outspoken in their protests at, for example, the efforts of the tobacco companies to push the medical researchers back into their laboratories to continue to collect superfluous evidence about a "possible" connection between smoking and lung cancer.

Mr Maxwell is an optimist, and in the last chapter of his admirable book he assures the reader that the revolution is already under way. In the past ten or fifteen years, as he points out, there has been a multitude of developments, within and without the academic world, that can be interpreted as attempts to implement the philosophy of wisdom — Rachel Carson's *Silent Spring* (1962), Barry Commoner's *Science and Survival* (1966), the Club of Rome's Report, *The Limits to Growth*, by Meadows *et al.*, Schumacher's *Small is Beautiful*, Higgins's *The Seventh Enemy* and many other seminal writings on the human implications of science. For those of us who practise science, the book's take-home message is that every scientist should ask himself, at every level from the most particular to the most general, exactly why he is doing what he is doing, and whether he would not be better advised to do it differently, or even do something entirely different. Is he content merely to acquire an expertise, which he can sell profitably in the market place; or should he find a vantage point from which he can make sure that his efforts, and those of his colleagues, serve, in the broadest possible sense, the interests of his fellow human beings? If *From Knowledge to Wisdom* stimulates a serious debate on such issues, it will have served a worthy purpose. □

Christopher Longuet-Higgins is Royal Society Research Professor in the Centre for Research on Perception and Cognition, University of Sussex.

Martha's revenge

David L. Hull

Biophilia: The Human Bond with Other Species.

By Edward O. Wilson.

Harvard University Press: 1984. Pp. 157. \$16.50, £13.20.

In this, his most recent book, E. O. Wilson presents three interrelated messages — that science provides a sort of knowledge that the humanities cannot produce; that genes are more important in making us what we are than many of us would like to believe; and that buried in our genetic make-up are a variety of propensities including a fascination for other living creatures and an unconscious longing for an environment such as the savannas in which we evolved. Wilson has pursued these themes before in a variety of guises — a general text, a philosophical treatise and, with C. J. Lumsden, dressed up in mathematical garb. In *Biophilia* he tries his hand at a more literary approach.

From its beginnings, science has found itself in conflict with not only theology but also the humanities, if for no other reason than theologians and humanists controlled the universities. Poets and literary figures complain that scientific knowledge destroys any genuine appreciation of nature; as Tennyson put this conviction, "Science grows and Beauty dwindles". Scientists, it is said, reduce the wonders of nature to a ceaseless war between selfish genes and explain humankind's higher faculties in terms of neurons and the hormone-secreting cells of the limbic system. Although scientists currently dominate the intellectual scene, defenders of the humanities retain some of their old feeling of superiority. Perhaps scientists can lay trans-Atlantic cables and unravel the molecular structure of the genetic material, but they can never hope to provide the higher sorts of understanding afforded by the humanities. Martha might well busy herself serving the guests, but Mary seated at her saviour's knee has the better part.

Wilson has written for his scientific audience several times over. In *Biophilia* he sets his sights on the great masses on the other side of C. P. Snow's great cultural divide, those readers who are more impressed by striking imagery than by scientific argument and masses of data. If humanists are the "shamans of the intellectual tribe, wise men who interpret knowledge and transmit the folklore, rituals, and sacred texts" (p.58), then Wilson will become a shaman. Here his style is consciously literary, a series of images designed to evoke the appropriate response in his reader. A peccary stolen from its natural habitat and tethered beneath the eaves of a rude hut becomes a "mute speaker trapped inside the unnatural clearing, like a

messenger to me from an unexplored world" (p.5). The mythical status of the serpent hints at the hold that our evolutionary past still has on us. The adventures which Wilson had as a young boy with snakes in the woods and pools around his home in the Panhandle of northern Florida illustrate the aversion and fascination which human beings have for snakes. His scientific description of the Bird of Paradise becomes a "metaphor of what humanists dislike most about science" (p.54).

For me at least, Wilson's attempts at this more literary style are effective. Everyone knows that the Sun is the ultimate source of energy for life on Earth. As Wilson puts it, "After the sun's energy is captured by the green plants, it flows through chains of organisms dendritically, like blood spreading from the arteries into networks of microscopic capillaries" (p.8). The most distinctive sound in the primary tropical forest, according to Wilson, is a "sharp crack like a rifle shot, followed by whoosh and a solid thump. Somewhere a large tree, weakened by age and rot and top heavy from layers of vines, has chosen that moment to fall and end decades or centuries of life" (p.27). To help his readers visualize the evolutionary process, Wilson contrasts various time frames. Everyone is familiar with organismic time: a leafcutter ant returns to its nest and drops the bit of leaf it has been carrying (a feat equivalent to a man running a four-minute mile while carrying 750 pounds). In biochemical time, organisms freeze in their tracks as a "nerve cell discharges: along the length of its membrane, the voltage drops as sodium ions flow inward" (p.42). At a level higher than the single organism, a leafcutter colony resembles a gigantic amoeba, a sort of super-organism. "Its foraging columns snake out like pseudopods to engulf and shred plants, while their stems pull the green pieces down holes into the fungus gardens" (p.36). In evolutionary time, even colonies lose their identity as they merge into an even more inclusive entity enlarging its territory, congealing into local populations, speciating and becoming extinct.

Wilson tries to give the humanities their just due. "The aim of art is not to show how or why an effect is produced (that would be science) but literally to produce it" (p.62). But with this one exception scientists can do everything that artists and writers can do and more. "Scientific innovation sometimes sounds like poetry, and I would claim that it is, at least in the earliest stages. The ideal scientist can be said to think like a poet, work like a clerk, and write like a journalist. The ideal poet thinks, works, and writes like a poet" (p.62). A poet can elicit images in us, but he "refuses to take us any further. If he goes on the precise image will melt into abstract descriptions; light and beauty will congeal into rows of formulas" (pp. 75-76).

Do I agree with Wilson's main messages?

How good are the arguments he presents? Has he made the right distinctions? To ask these questions is to miss the intent of the book. It is to elicit images and appropriate emotional responses. The clearing of jungles in the Amazon to make room for bony white cattle "can be defended (with difficulty) on economic grounds, but it is like burning a Renaissance painting to cook dinner" (p.25). This is not an argument, but it does the job as well or, for most people, better. In *Biophilia* Wilson uses the traditional techniques of the humanities to get his messages across, and one of them is that the humanities are not good enough, that the truths afforded by science transcend those of the humanities. In this book Martha gets her revenge. □

David L. Hull is Professor of Philosophy at Northwestern University, Evanston, Illinois.

Ecosystems unlocked

Mark Williamson

Key Environments.

General editor John Treherne.

Sahara Desert. Edited by J.L. Cloudsley-Thompson. Pp.348. **Galapagos.** Edited

by R. Perry. Pp.321. **Madagascar.** Edited

by A. Jolly, P. Oberlé and E.R.

Albignac. Pp.239.

Pergamon: 1984. Each volume £14.95, \$23.95.

WHAT is a "key" environment? For this series it is taken to be one of international ecological importance which is now, or soon will be, under threat of disturbance or destruction. Eight volumes have been advertised so far, and three are to hand; the others promised deal with Amazonia, Antarctica and Malaysia, and there will be two marine ones, on the Red Sea and the Western Mediterranean.

Is it possible to "provide specialists, as well as those who have an interest in the conservation of nature as a whole, with the essential facts [for] practical and effective conservation action" in books of 300 pages written by 20 or so authors? Not really: what in fact has been achieved are reasonably comprehensive and on the whole stimulating accounts of standard natural history, that is, geology and climate, higher plants, vertebrates. The coverage of conservation or management as such is meagre, and there is certainly not enough information to build, say, even a crude model of the effects of different strategies. Rather, *Key Environments* is "a scientifically accurate, concise and well-illustrated series of accounts" which does indeed "summarise the present knowledge of the flora and fauna".

The format is uniform, 18 x 24 cm, well set out and structured, and with black-and-white photographs or line drawings to nearly every page. The reproduction of

many of the photographs is disappointingly murky and dark, however, and although there are quite a lot of maps, the editors on the whole seem to think basic geography uninteresting, or maybe so familiar as not to be worth much mention in text or index. The series certainly deserves a wide readership amongst professionals, but particularly professionals unfamiliar with the areas concerned, and also among the wider audience presumably attracted by the Duke of Edinburgh's name (and no one else's) on the cover of each of the books.

Perhaps because of my interest in islands, I learnt most from the volume on the Sahara Desert. The editor is moderately successful in getting his authors to stay within the 100 mm isohyet, shown roughly on p.22 and more exactly on pp.106 and 329. The major question over this enormous area of $9 \times 10^6 \text{ km}^2$ (about the size of the continental United States) is whether the desert is spreading south into the Sahel. The editor's introductory chapter, Warren on problems of desertification, Allan on oases and some of the biological chapters discuss the matter, and in particular the problem of detecting permanent changes in areas of most erratic rainfall. Williams on geology, Wickens on flora and Milburn on archaeology and prehistory put the problem in an historical perspective. Only Smith on climate seems to doubt that there is any real change or real problem.

For me, the most surprising information

was about the waters of the Sahara: not just the temporary waters after the rare rain, or the old and modern, engineered, oases, but the natural habitats that supported animals such as crocodiles in the Central Sudan until this century. Again, though, the publishers might have been more careful with the supporting material — the occurrence of crocodiles in the Oued Ihmirou in the mountains of the Tassili N'Ajjer (southern Algeria) is recorded on p.214, but not indexed under either species or place; the mountains are shown on maps on pp.327 and 332, the Oued nowhere (nor in the *Times Atlas*; and can you define an Oued? Would it help to spell it Wadi?).

The Galapagos volume is, not surprisingly, the most successful, as a research station flourishes there and many research workers visit the islands. Physical oceanography, sea birds, seals, inshore fishes and the protection of the marine environment are all covered, as well as the terrestrial aspects. Conservation is dealt with quite thoroughly in Hamann's chapter, "Threats to the Vegetation", and in Hoeck's contribution on the introduced fauna — cattle, horses, pigs, dogs and, as so often, cats, rats and goats are all found on the Galapagos, and there has been some success in containing some of them. But the two chapters on the "Path of Conservation" are very threadbare. Comprehensive this volume is not. The only invertebrate mentioned is the introduced fire ant, herons and such birds come in only casually, while the chapter on terrestrial plants is a scrapbook, including an update of the standard flora and the odd suggestion that there are no endemic plants in the British Isles. Sulloway's mistaken claim about Darwin's finches and the *Origin of Species* reappears. Nevertheless this volume contains a lot of useful information, much of it not available elsewhere in such accessible form.

Many of the authors of *Madagascar* are French or Malagasy, but the English is consistently good. The emphasis is on the vertebrates, especially mammals to which four chapters are devoted. The accounts of economics and conservation, and of nature reserves and nature conservation are much to the point. So this volume is in its way even better than that on the Galapagos, but the content is distinctly thinner.

Threats to natural environments usually come from man's activities, among them the introduction of alien species. Understanding the problems of natural ecosystems requires as much study of man's influence as of the ecosystems themselves, and to that extent the series is lacking. Perhaps later volumes will put this right. The three published to date are nevertheless outstanding accounts of important ecosystems. □

Mark Williamson is Professor of Biology at the University of York. He is author of *Island Populations* (Oxford University Press, 1981), which appeared in paperback last year.

Is it getting better all the time?

Lynton K. Caldwell

The Resourceful Earth: A Response to Global 2000.

Edited by Julian L. Simon and Herman Kahn.

Basil Blackwell: 1984. Pp.585. £14.95, \$19.95.

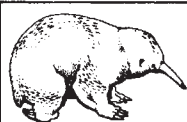
IN ANCIENT times he who brought bad news to the king risked being put to death for having offended royal dignity and raised doubt about the security of the realm. Today, bad news is no more welcome; *The Resourceful Earth* has been published in the belief that "the world is ready to turn its back on pessimism and is waiting to hear some good news". But in order to secure the plausibility of the "true good news", the "false bad news" must be exposed and refuted. The bringer of bad tidings offensive to the editors and sponsors of *The Resourceful Earth* was the *Global 2000 Report to the President* (1980), commissioned by President Jimmy Carter and released during his last few months in office. Herman Kahn and Julian Simon are the self-appointed executioners of that messenger of "gloom and doom", whose forecasts of bad times to come extend beyond the United States to the entire planet Earth.

The publishers declare that *The Resourceful Earth* challenges the conclusions of *Global 2000*, and presents scientific evidence that sets the record straight. The Executive Summary is more explicit:

The original 1980 *Global 2000 Report to the President...* is frightening. It received extraordinarily wide circulation, and it has influenced crucial governmental policies. But it is dead wrong. Now *The Resourceful Earth*, a response to *Global 2000*, presents the relevant reliable trend evidence which mainly reassures rather than frightens.

The aims of the book are thus twofold: first, to refute the "bad science" of *Global 2000* and, second, to reassure a public frightened and confused by that report's prophecies of global disaster. In achieving its objectives, *The Resourceful Earth* may be regarded as a one-quarter success and a three-quarters failure, a judgement which calls for explanation. The greater part of the book is not the work of Kahn or Simon; it consists of invited papers (some previously published), contributed by more than 20 authors of good repute in science and economics. Its chapters are heavily laden with statistical facts, many purporting to show that "aggregate global and U.S. trends are improving rather than deteriorating".

The failure of *The Resourceful Earth* is not primarily in the 21 contributed chapters, some of which do not actually support the editors' thesis (for example



Quaternary Extinctions

A Prehistoric Revolution

Paul S. Martin & Richard G. Klein

More than forty contributors present evidence for major theories—climate vs. overkill—about the disappearance of mammals and birds toward the end of the last ice age. 900 pages (7 x 10), illus. \$65.00 U.S., August 1984.

Darwin: Structure and Distribution of Coral Reefs

foreword by Michael Ghiselin
256 pages, \$7.95 paper, 1984.

People of the Desert and Sea



Ethnobotany of the Seri Indians

Richard Felger & Mary Beck Moser

Records a contemporary hunter-gatherer society's uses of more than 400 plants found in the Sonoran Desert and Gulf of California. About 450 pages (8 1/2 x 10 1/2) with more than 300 illus. \$65.00 U.S., December 1984.

University of Arizona Press

1615 E. Speedway, Tucson, AZ 85719
Member of the Eurospan Group

those by Roger Revelle and Gilbert F. White), and many of which do not undertake explicitly to refute *Global 2000*. Rather it lies in the selective uses and interpretation that the editors make of data pertaining to the environmental future, and in the extravagant claims they and their publisher make as to the achievements of the book:

This is the famous report which demolishes the *Global 2000 Report to the President*. Published here for the first time, it is both a devastating indictment of all doomsday books and also the most scientific inquiry into the future ever organized.

That claim to fame is at least questionable: famous to whom and for how long? The book may have acquired instant celebrity status among those who, aware of its genesis, welcomed at last a refutation of "wrongheaded environmentalism". On the other hand, it may have been regarded as famous — or rather infamous — by those who see its objectives as harmful and its findings distorted and tendentious. Nonetheless except as promoted by believers in its gospel, it seems improbable that *The Resourceful Earth* will achieve the notoriety of *Global 2000* — a point that its editors concede. So its principal claim to fame should be that it demolishes *Global 2000*; in my opinion it has not done so.

While specific findings and predictions of *Global 2000* may be questioned, it is hardly possible to demonstrate that the report as a whole is "dead wrong". Its general validity — and the thesis of the editors of *The Resourceful Earth* — can only be established by events in the future. But *Global 2000* is supported by a much broader and diverse base of evidence than is *The Resourceful Earth* because the two reports differ fundamentally in their approach to evidence. *Global 2000* sought all available reliable data to indicate the trends that, if unmodified, would determine the state of the world by the year 2000 and thereafter. The projections of *Global 2000* followed from evidence unbiased by an intended outcome; the contributors did not set out to find and demonstrate that "the world in 2000 will be more crowded, more polluted, less stable ecologically, and more vulnerable to disruption than the world we live in now". Rather their conclusions reflected the evidence as they read it, which corresponded to the findings of a series of comprehensive analytical models and simulations of the future undertaken by various investigators under various sponsorships since the publication of *The Limits to Growth* report in 1972. *The Resourceful Earth*, on the other hand, was structured specifically to attack *Global 2000*, and the selection and use of evidence appears to have been restricted to this end. The editors make no claim to summarize all the relevant evidence regarding environmental trends — their opening statement in the Executive Summary declares that the book presents "the relevant reliable trend evidence which mainly

reassures rather than frightens".

The Resourceful Earth is advertised as "a devastating indictment of all doomsday books". But "doom", meaning death, total ruin or final judgement, is not the prediction of the greater part of the literature



British Tourist Authority

Living with pollution — "if there is reason to believe that a large part of the evidence of environmental stress and deterioration is valid, does not a responsible government have an obligation to bring this evidence to public notice?"

which Kahn and Simon affect to indict. Major disasters have been forecast, or more often conditionally conjectured, by various writers including well-informed scientists and committees of the International Council of Scientific Unions. The truly "doomsday" books are those dealing with nuclear wars — which Herman Kahn believed were also unfounded and alarmist. If there is reason to believe that a large part of the evidence of environmental stress and deterioration is valid, does not responsible government have an obligation to bring this evidence to public notice?

Kahn and Simon do not think so. "We believe", they write, "that the government should not take steps to make the public more 'aware' of issues concerning resources, environment and population", explaining that "the public has been badly

served by having been scared by a very large volume of unfounded and/or exaggerated warnings about these matters". But what proportion of the public has been "scared" by *Global 2000* and similar reports is not made clear. Evidence of a trend towards greater environmental concern in the United States, Britain and Western Europe has been documented by William Cameron Mitchell of Resources for the Future, and by Lester W. Milbrath of the State University of New York. These findings indicate that concern is greatest among the most educated (and presumably best-informed) sectors of the population. Is it plausible to believe that the presumably best-educated and most fully informed members of society can be more easily stampeded into hysteria by misleading pseudo-science than the under-informed members of the body politic?

It is, in fact, far from sure that *Global 2000* is the principal culprit in upsetting the public's emotional equilibrium. In his essay on "Global Climate Trends" in *The Resourceful Earth*, H.E. Landsberg concludes that "although much of the climate portion of *Global 2000* and its sequel is quite sound, the conclusions drawn from them in the news media picked only the frightening aspects up for communication to the public". A harsher indictment of the media is offered by Bernard L. Cohen in his "Statement of Dissent" in which he declares that:

Our government's science and technology policy is now guided by uninformed and emotion-driven public opinion rather than by sound scientific advice. Unfortunately, this public opinion is controlled by the media, a group of scientific illiterates drunk with power, heavily influenced by irrelevant political ideologies, and so misguided as to believe that they are more capable than the scientific community of making scientific judgements.

Global 2000 is also not faulted for scaring the public by the authors of a chapter on "Global Trends in Nonfuel Minerals". Instead they declare that "The merit of *Global 2000* with respect to nonfuel mineral(s) is that it has not said much that is wrong. The defect is that it has not said anything that was not well-known, and it failed to identify and discuss the major policy issues" — which is just what Kahn and Simon say the government should not do. Yet readers of *The Resourceful Earth* might find disconcerting the opinions of two contributors on the prospective use of coal. In Chapter 15B, Petr Beckmann states that "The world's use of coal will increase to some 5 billion tons by the year 2000, in part due to the pressure to replace oil, in part due to the growing demand in developing countries" (i.e. China). Yet the concluding sentence in Chapter 20, "The Hazards of Nuclear Power" reads: "Every time a coal-burning plant is built instead of a nuclear plant, many hundreds of people are condemned to premature death". Is this "reassuring"?

The final claim for *The Resourceful*

Earth is that it is "the most scientific inquiry into the future ever organized". This seems unsupportable, if only because it is doubtful whether any estimate of the future thus far published can be shown to be "the most scientific". Is this a creditable boast for a publication never subjected to peer review? I would contend that none of the projections of the future made thus far are "scientific" except and in so far as they are based upon the best available scientific evidence. Any projection of the future is fraught with uncertainty and the possibility of error — neither *Global 2000* nor *The Resourceful Earth* are exceptions.

Allowing a limited validity to the aphorism "a little knowledge is dangerous", a little *valid* knowledge today may be greatly preferable to common-sense ignorance. The extraordinary growth of scientific knowledge during the past quarter-century provides more and better data than that heretofore available for assessing future trends. But the data do not automatically provide this assessment — neither do the science specialists who develop the data. Individual scientists, including contributors to *The Resourceful Earth*, are prepared to set forth the findings of their specialties but are not necessarily prepared or able to relate them to findings in other disciplines. The global models that Kahn and Simon disparage are attempts to develop syntheses among diverse but related scientific findings — an enterprise hardly possible before the development of powerful and sophisticated computers. Whether this modelling is "science" or merely technique may be a matter of definition; whatever it may be called, it is a relatively recent development. Yet even with its inevitable imperfections (which only experience can correct) it is surely more reliable than Kahn and Simon's simplistic ruler-and-pencil trend analysis.

The foregoing observations indicate where I believe *The Resourceful Earth* has failed to measure up to its billing. But what of its one-quarter success? To me, its limited merit is the occasion which it provides for a careful, critical examination of the conclusions that it purports to refute, and the data it has assembled in the respective contributions on a range of environmental issues. I regard the assumptions of its editors as wrong, but nonetheless deserving of open-minded examination. The Kahns and Simons of the environmental movement are for their part not without comparable sin. In responding to the allegations of *The Resourceful Earth*, the environmentalists may discover shortcomings in their own analysis and be pressed to state more precisely why *Global 2000* and its associated literature deserve to be taken seriously. □

Lynton K. Caldwell is Bentley Professor of Political Science at Indiana University. His most recent book is *International Environmental Policy: Resources and Dimensions* (Duke University Press, 1984).

Making war on the world

Alastair Hay

Environmental Warfare: A Technical, Legal & Policy Appraisal.

Edited by Arthur H. Westing.

Taylor & Francis: 1984. Pp. 107. £12, \$21.

Herbicides in War: The Long-Term Ecological and Human Consequences.

Edited by Arthur H. Westing.

Taylor & Francis: 1984. Pp. 210. £15, \$33.

SEEDING clouds to increase rainfall, breaching reservoirs to cause flooding and denuding forests with herbicides are some of the better known techniques of environmental warfare. All three activities were carried out by US armed forces in Vietnam. Cloud seeding was a failure. Dykes which were breached on occasions caused some flooding, but this never seriously undermined North Vietnam's war effort. It was the defoliation programme which, as far as the US military was concerned, could be pronounced successful. Large tracts of naked forest provided no cover for the enemy and discouraged all movement in the area.

In war the object is to win. If this means destroying buildings or countryside, then history is replete with examples of military commanders who have not thought twice about razing an area; for the soldiers in the field it is a legitimate activity. But as wars become ever more destructive, and as awareness grows of the fragility of our environment, we must ask whether this form of warfare should be allowed to go unchecked.

Environmental Warfare and *Herbicides in War* are both edited volumes, emanating from the Stockholm International Peace Research Institute, which address these issues. If they have one message in common, it is that such practices cannot continue. The first book discusses the subject from a global perspective, whereas the second is concerned exclusively with assessing the impact of the military use of defoliants in Vietnam from 1962 to 1971.

Indochina has not been the only testing ground for some of these refinements of warfare, merely the latest. Dykes were breached in the Franco-Dutch war of 1672–1678 to halt the progress of attacking French forces, and with some success. But the single most devastating act of destruction recorded was during the Sino-Japanese war of 1937–1945, when, in June 1938, Chinese forces dynamited the Huayuankow dyke on the Yellow River. Several thousand Japanese soldiers drowned in the subsequent flood and the advance of the Emperor's forces on this front was stopped. The Japanese, however, were not the only victims. Hundreds

of thousands of Chinese were also drowned, millions of acres of crops were destroyed and life on the banks of the Yellow River remained unsettled until 1947 when the river was once again brought under control. It is this latter point that is at the crux of any discussion of environmental warfare — the effects are indiscriminate. Non-combatants invariably bear a disproportionate number of the

IMAGE
UNAVAILABLE FOR
COPYRIGHT
REASONS

Chemical blitz — mangrove forest in South Vietnam before (top) and after spraying with herbicide.

casualties, and in Vietnam too it was the civilians that were the real losers.

Herbicides in War is the published proceedings of a conference held in Ho Chi Minh City (formerly Saigon) in January 1983. The meeting was called to assess the long-term effects of the defoliation programme and, as the book makes clear, those effects were devastating. During the spraying programmes, some 10.3% of Vietnam's inland forests, 36% of mangrove forests, 3% of cultivated land and 5% of other land was sprayed one or more times with a variety of "anti-plant agents". Some 19 million gallons of defoliants were used, the most common of them being Agents Blue, Orange and White — so-called because of the painted bands on the drums in which the herbicides were shipped to Vietnam.

The book contains the reports of working groups which evaluated the impact of the defoliants on Vietnam's ecology and on the health of the population, and the ways in which the toxic contaminant dioxin (2,3,7,8-tetrachlorodibenzodioxin) present in Agent Orange can be measured in the environment. The background to these subjects is provided, in the main, by

Vietnamese scientists, while the overall findings are set in context by general overviews written by American, British and Swedish scientists. Their conclusions are that Vietnam's inland and mangrove forests not only suffered immediate and extensive damage as a result of the spraying, but that the effects persist. Only a replanting programme — already begun by the Vietnamese — on a massive scale will heal the scars of war. For this work to succeed time and international help will be required.

The effects of the various chemical agents on the health of the population is more difficult to judge. Vietnamese scientists believe that the herbicides caused an increase in the incidence of liver cancer and in the number of children born with handicaps. Regrettably it is too early to assess both claims; the evidence is still far from complete and much more information will need to be gathered to resolve the issue. Meanwhile the population of Vietnam has to live with uncertainty. It will be little consolation that the havoc visited upon their country has led to moves to prevent the same thing happening again elsewhere.

In 1977, outrage over the attempted manipulation of the environment of

Indochina went a stage further than mere condemnation, with the enactment of the Environmental Modification (ENMOD) Convention. The Convention was designed to prohibit acts of war intended to change the natural environment, but fell short of what was required. *Environmental Warfare* reviews the agreement and points out how the text of the 1977 drafting should be tightened to exclude totally any manipulation of the environment by armed forces (some is still permitted) and to outlaw research on methods of manipulation for hostile purposes.

In the heat of battle there will be no time, nor desire, to agree codes of conduct; such agreements need to be hammered out in calmer times. A nuclear war, with the ensuing nuclear winter, may render some of these resolutions academic, but in the event of more conventional warfare we will be glad of them. These two publications will be invaluable for bringing that lesson home. □

Alastair Hay is a Lecturer in the Department of Chemical Pathology at the University of Leeds. With Seán Murphy and Steven Rose, he is co-author of No Fire, No Thunder: The Threat of Chemical and Biological Weapons (Pluto Press, 1984).

Regulation damage

Anne H. Ehrlich

Environmental Policy in the 1980s: Reagan's New Agenda.

Edited by Norman J. Vig and Michael E. Kraft.

Congressional Quarterly Inc., 1414 22nd Street NW, Washington DC 20037: 1984. Pp.377. Pbk \$12.95.

SINCE shortly after Ronald Reagan took office as President of the United States, environmentalists have protested that violence is being done to America's hard-won federal regulatory apparatus for environmental protection and resource management. The means, it is alleged, are administrative fiat and selective budget-slashing. Now a group of scholars with an impressive array of academic credentials has analysed the Reagan team's record in administering the government's resource and environmental programmes. They have produced a detailed indictment of that record, as resounding as anything so far issued by even the most militant of environmental organizations.

Some of the territory explored by Norman Vig, Michael Kraft and the 15 other contributors to this collection is familiar to the public. Most notably there have been the scandalous goings-on at the Environmental Protection Agency (EPA)—the reign and departure in disgrace of its administrator Anne Gorsuch-Burford, and the belated attempt to rescue the agency's remaining shreds of

credibility by appointing William Ruckelshaus as its head in 1983 — not to mention the achievements of Reagan's remarkably insensitive (both socially and environmentally) first Secretary of the Interior, James G. Watt.

Meanwhile, however, other events of questionable legitimacy have occurred without attracting the limelight — the dismantling of the Department of Energy's information-gathering and research capabilities and its programmes to encourage conservation and development of alternative energy sources; the mutually contradictory policies of forest management in the Department of Agriculture; and the abdication of the nation's role of world leadership in environmental protection and global resource management. An example of the last, which did attract public attention but occurred too late for inclusion in this book, was the Reagan administration's ludicrous performance at the United Nations' World Population Conference last August in Mexico City.

Vig, Kraft and their co-authors scrutinize these and other actions, the effect (and, evidently, intent) of which has been to weaken, if not cripple, the environmental regulatory power of the federal government. The book also presents a history of environmental legislation and regulation at the federal level, a discussion of the misconstrued electoral "mandate" for the sweeping changes that were effected, and an assessment of the long-term damage inflicted on the government's institutions by the Reagan regime in just three years.

The administration's attempts to revise

Advertisements

If you want it tomorrow...
call us today
01-455 9823

COMPUTER CALCULATORS PLOTTERS

HP 11C (21C Mem Sci simline)	£63.00
HP 15C (Adv Sci c/ mem) 67 storage reg.	£84.00
HP 41CV (Alphanumeric sci. computer)	£164.95
HP 41C Card reader	£129.95
HP 41C Printer (82 143A)	252.00
HP 41CX & HP 71 Printer (82162A)	£295.00
HP 16C (Hex Converter)	£94.00
HP 71 (New Computer)	£410.00
HP 41CX (New version of CV) with ext. functions/ timer	£210.00
Complete range of HP accessories in stock	

HP BUDGET PACKS

HP 41CV together with card reader and printer 82143A. All complete to manufacturers spec. **£520.00**

Pixy Plotter (3 pen A4 8 colours)	£399.00
HP 7470A (A4 2 Pen Plotter)	£799.00
HP 7475 (A3 6 Pen Plotter)	£1399.00
(A3 Pen Plotter)	
SHARP (Graphics Printer/ Calc (4 colour)	£137.95
SHARP 1500A Pkt. Comp. 8K Mem	£147.50
CE 158 RS232 and Cent/1	£130.00
CE 150 printer cass./f	£130.00
CE 152 Cassette	£36.95
CE 158 BK (Add on mem with own bat.)	£79.00
PC 1251 (Wallet size Computer)	£69.50
PC 1401 (Pocket Computer) sci. cal.	£59.95
CE 125 (Printer/ Tap Rec. for above	£86.50
CE 126 (Cassette/ Printer) for above	£59.95
Epson HX20 Briefcase Computer	£375.00
Epson PX8 (Portable computer)	£775.00
Epson LQ 1500 (New 200 CPS printer)	£885.00
Casio PB 700 (H/Held Computer exp to 16K)	£127.00
Casio PB300 (Computer 2K with built-in printer)	£89.95
Casio FX 802P (222 Mems built-in printer)	£79.95

**GOODS FULLY GUARANTEED
PRICES EXCLUDE VAT AND P&P**
Company, hospital and Government orders accepted by phone
Barclaycard: Access/Visa accepted by phone

EXPORT ORDERS WELCOMED

MOUNTAINDENE

22 COWPER ST. LONDON EC2

ITE publications

METALS IN ANIMALS

Editor D. Osborn

ITE Symposium volume no. 12 covering many aspects from effects of metals on wildlife to advanced chemical techniques (NMR).

1984 0 904282 77 5 paper 77pp A4 £3.50

BRITAIN'S RAILWAY VEGETATION

Caroline Sargent

Habitat, vegetation forms and influences that have combined to produce them are described; decline and success of railway species; predictive model and conservation strategy.

1984 0 904282 76 7 paper 34pp A4

MOORLAND MANAGEMENT: a study of Exmoor

G.R. Miller, J. Miles & O.W. Heal

Ecological management for agriculture, sport and conservation, with special reference to Exmoor.

1984 0 904282 78 1 paper 118pp A4 £4.50

Three of the recent publications from the Institute of Terrestrial Ecology. Full list of all publications available.

ITE, Administrative Headquarters, Monks Wood, Abbots Ripton, HUNTINGDON, Cambs. PE17 2LS.

environmental laws, the conventional strategy by which presidents make policy changes, were largely rebuffed by Congress, which in turn had been subjected to increasing pressure from the public, both directly and through revitalized environmental groups. But Congress has little control over administrative decisions; while it could partially restore ravaged budgets, it could not force administrators to spend the money.

To a large extent, as the authors make clear, the administration's actions have been counterproductive to their intended ends. The fiasco at the EPA not only defeated the goal, once shared by regulated industries and environmentalists alike, of reforming the existing regulatory framework to make it more effective, efficient and fair; it backfired. By politicizing regulatory affairs, and either firing or alienating the EPA's professional staff, the EPA administration under Gorsuch set back the clock on much needed regulatory reassessment. Richard Andrews, author of the chapter on the EPA, comments:

The attempt to achieve radical policy change by means of ideological loyalist appointees untainted by experience thus stands as probably the single most reckless tactic, and clearest failure, in Reagan's policy strategy.

Similarly, the administration's efforts to weaken the Clean Air Act may have led to more, not less, rigidity in regulation and increased mistrust between regulators and the regulated.

The Reagan team's actions have also been self-contradictory. Complaining on one hand that insufficient information existed to justify certain programmes, it proceeded to kill research projects that could provide the necessary data. Describing the relinquishment of the Department of

Energy's responsibilities for data-collecting and analysis, Regina Axelrod drily observes: "No government in recent times has sought to beat its opponents by declaring war on information".

Professors Vig and Kraft have produced not only a stinging critique of Ronald Reagan's environmental policies, but a useful, general text on the making and implementing of environmental policy at the federal level in the United States. Its viewpoint is generally pro-environment, yet it is sympathetic to the needs of business and industry for reasonable, consistent and even-handed environmental regulations. The book unfortunately suffers from a certain amount of repetition and overlap, but that perhaps is an inevitable result of the difficulty in orchestrating the efforts of 17 authors.

Beyond its value as a text on environmental policies, the book provides an instructive case-history of an "administrative presidency" — a presidency that effects policy changes through administrative decisions. In this instance, the changes have been both unprecedentedly radical and lacking public consensus. Political scientists no doubt will find food for thought for years to come, not only in the lessons the Reagan presidency offers for eluding the complex checks and balances built into the government's structure and the degree of change that it successfully brought about despite them, but also in the failure of the American public to hold Ronald Reagan personally responsible for policies that it clearly does not support. □

Anne H. Ehrlich is a Senior Research Associate in the Department of Biological Sciences, Stanford University.

Boffin's progress

Robert Cockburn

Bernard Lovell: A Biography.

By Dudley Saward.

Robert Hale, London: 1984. Pp.320. £12.95.

BORN in Gloucestershire in 1913, Sir Bernard Lovell is a worthy successor to the bold buccaneers who thrived in the West Country during the reign of the first Elizabeth. He is a man of ruthless energy, an outstanding scientist, an engineer *faute de mieux*, but above all an inspired leader.

The author of this biography was a Group Captain on the staff of Bomber Command during the Second World War, collaborating closely with Lovell, and he has remained an intimate friend of the family ever since. So Saward is well qualified to write a definitive account of Lovell's life and works, though he sometimes tends to include facts from his own experience which are not particularly rele-

vant to the story as a whole; for instance there is perhaps rather too much detail on Bomber Command operations in Chapter 7, and on the activities of the Renwick group on behalf of a space launcher in Chapter 18. But otherwise this is a careful, sympathetic and well-researched book.

The first part deals with Lovell's time at the Telecommunications Research Establishment (TRE) during the War, where he was project leader for the air-to-ground radar, H₂S. This was a quite unique establishment that had brought together most of the best young scientists in Britain to develop the new techniques of radar. There was hardly any application which did not meet some desperate demand from the Services, and backed by over-riding priority status the pace of development was tremendous. Henry Tizard had established a mutual trust between the airman and the boffin, and this intimacy was maintained and expanded throughout the War. The speed of response to the changing operational requirements which this permitted was a major contribution to victory.

Bomber Command had been completely

unprepared for night bombing and its performance over Germany was lamentable. The introduction of H₂S, first into the Pathfinders and then to the main bomber force, improved its efficiency by an order of magnitude. It was to this end that Lovell directed his redoubtable energies; he transgressed every rule in the book, scrounged endlessly, and by sheer bloody-mindedness cut through all the usual organizational restraints. It was a personal achievement of the highest order. As the end of the War came in sight, and the pressures on TRE eased, many of the leading figures collapsed after four years of unremitting toil. Lovell was no exception, although he recovered rapidly.

As soon as he arrived back at the University of Manchester in 1945, Lovell applied himself with characteristic determination to equipping the green-field site at Jodrell Bank with apparatus purloined from the Services, and to laying the foundations for his subsequent career as a radioastronomer. It was in 1950 that the idea for a huge steerable telescope caught his imagination, and the fruitful collaboration with H.C. Husband commenced — Husband was the only engineer in the country prepared to design such a novel and advanced structure. Nowadays we are inured to the inevitable escalation of costs during the development of an advanced project. At that time the Department of Scientific and Industrial Research, the funding agency, was unused to such expensive investments, and throughout the seven years that it took to bring the telescope into operation there was continuous wrangling over finances. Used to the lavish wartime expenditure, Lovell found peacetime restraints irritating in the extreme; but loyally backed by Manchester University, and by P.M.S. Blackett and other far-sighted senior scientists, he forced the project through with the same drive that he had used with such effect during the War. There were undoubtedly frequent bouts of depression and frustration, but his resolve remained firm.

At the very last stage, when the equipment was on the verge of operation, public opinion swung in his favour when it was realized that the telescope was the only facility in the world capable of tracking the Sputniks. For the next three years Jodrell Bank was used by both the Americans and the Russians as a vital tracking station for their space probes. Lovell was by now an international figure, and was invited to the United States and USSR where he was received with the greatest courtesy and respect, and was able to arrange several collaborative projects.

In addition to all his research work Lovell found time to write a large number of books and give lectures, and for numerous public duties. As a member of the Science Research Council and chairman of the Astronomy Space and Radio Board, he became involved in wider issues of national policy. At one point, in

the late 1960s, he generously supported Ryle's synthesized aerial at Cambridge even though this meant abandoning his own cherished ambition for a thousand-foot telescope. But despite his involvement in so many national issues and recurring budget cuts, Lovell continued over the years to update and extend the facilities at Jodrell Bank, culminating in the multi-telescope interferometer which has vastly improved the resolution available. Manchester University now has a research centre which will keep it in the forefront of radioastronomy for many years, an



Central Office of Information

Man and machine — Bernard Lovell at Jodrell Bank, 1977, with the telescope in the background.

impressive monument to one man's ability, ingenuity and dogged determination.

Apart from giving his readers a portrait of Lovell, in this book Saward provides a valuable insight into the efforts necessary nowadays to bring about a major scientific advance. It is no longer sufficient to convince one's peers of the scientific merits of a particular project. These have to be explained in terms which a layman can appreciate. Allies have to be courted, doubters converted and critics confounded. Expenditure must be justified against competing projects and the campaign will need to be sustained for many years.

The biography concludes with Lovell's retirement in 1980 as Professor of Radioastronomy at Manchester. But this cannot be the end of the story — it somehow seems unlikely that he will be content to pass his time cultivating the garden, playing the piano and watching cricket, his leisure interests listed in *Who's Who*. □

From 1939 to 1945 Sir Robert Cockburn worked on radar at the Telecommunications Research Establishment, Malvern. He was later Chief Scientist of the Ministry of Aviation and Director of the Royal Aircraft Establishment, Farnborough.

The bomb and the bumble-bees

Walter Gratzer

Late Night Thoughts.*

By Lewis Thomas.

Oxford University Press: 1984. Pp. 175. £9.95.

"Is ditchwater dull? Naturalists with microscopes have told me it teems with quiet fun." Thus G.K. Chesterton, and his words could serve as a text for some of Lewis Thomas's best essays. This, his third collection, will not disappoint his devotees, who will find here the familiar blend of reminiscence and reflection, the excursions into etymology, and withal an artful simplicity of style, illuminated by periodic shafts of startling insight. Thomas descants on health care, the funding of science and the nuclear arms race; he ruminates on the march of molecular biology, on natural history and the benefits of biomedical research. One of the threads that runs through his discourse is his entrancement with the intricacies of nature; Thomas sees the Earth as a kind of exquisitely crafted, bejewelled Fabergé egg,

the loveliest object afloat round the sun, enclosed in its own blue bubble of atmosphere, manufacturing and breathing its own oxygen, fixing its own nitrogen from the air into the soil, generating its own weather at the surface of its rain forests, constructing its own carapace from living parts.

There is touch of mysticism here, as also elsewhere, for example in his fascination with the mind of the bee.

The second thread that runs through the book is a bracing optimism about science, present and future — the only game in town, Thomas calls it — with perhaps even a hint of *hubris*; the problems engendered by science, he believes, will be put right by scientists, iatrogenic diseases cured by doctors. He even has a good word for the embattled discipline of sociology, which he expects soon to see emerge as an exact science.

These robust views contrast oddly with his dark perception of a future with the Bomb. Brooding on the unthinkable, he gives vent to the following: "If I were very young, sixteen or seventeen years old, I think I would begin, perhaps very slowly and imperceptibly, to go crazy". Thomas is not, to my mind, at his best in this rhetorical vein; what one looks for from a writer and scientist of his stature is something more reasoned than an evocation of the horrors of Hiroshima: Freeman Dyson, in his admirable series of articles in the *New Yorker* (now a book), showed that it can be

done with no loss of impact. But as that eupeptic divine, Sydney Smith remarked about his own meditations on the human condition, cheerfulness will keep breaking in; Thomas proposes, as a safeguard against nuclear war, huge armies of conscript hostages, American and Russian, to be kept in perpetual motion across each others' countries, lodged in luxurious Pullman trains, full of beer, running to undisclosed timetables. Such travel would both broaden the minds of the soldiery and restore the railways to solvency, while ensuring that no missile ever leaves its launching pad.

The book abounds in such *jeux d'esprit*. Thomas's imagination is caught, for example, by the biology of the lie-detector: if the telling of a lie causes sufficient anxiety that it should reveal itself in a measurable physiological change, he wonders, are there perhaps other manifestations, such as secretion of a peptide or some other product from the skin, that could be detected and quantitated by microanalytical techniques? And is there then a genetic basis for the ethical imperative, a biochemical morality? Dr Thomas explores with some relish the social consequences that such a discovery might bring in its train. (Ambrose Bierce, on the other hand, defines conscience as a small, still voice that warns you that someone may be watching.)

Many of these essays will serve splendidly for starting dinner-table conversations; the "Seven Wonders of the Modern World", for instance, should be a diverting game to play (even though Dr Thomas's first nomination — the bacterium that was supposed to thrive at 250°C under pressure in thermal vents under the Pacific Ocean — has already been exposed for an illusion, due probably to contamination by a less aristocratic organism).

Thomas's purpose is also more serious, however, for he does not address himself only or even primarily to scientists. I do not believe there is anyone writing today who better represents to the outside world how basic and applied research, and indeed scientists, function. He makes clear for example the essential unpredictability of the process of discovery. It was the organic chemist, Homer Adkins, who defined basic research as shooting an arrow into the air, and where it lands painting the target. The implications of this truth have urgently to be brought home to our paymasters, the public, and their none-too-obedient servants, the politicians. Let us hope *Late Night Thoughts* falls into the hands of some of the more literate of them and that some of the lessons penetrate their preconceptions. The rest of us will settle for the continuing stimulation and entertainment that Dr Thomas so gracefully affords. □

*In the United States *Late Night Thoughts on Listening to Mahler's Ninth Symphony*, published by Viking, \$12.95.

Walter Gratzer is in the Medical Research Council's Cell Biophysics Unit, King's College, University of London.

Glittering prizes

Joseph Silk

A Hundred Billion Stars.

By Mario Rigutti. Translated by Mirella Giacconi.

MIT Press: 1984. Pp. 285. \$25, £26.15.

TO ALL Earth-bound observers, stars are brilliant points of light, unresolved even by our largest telescopes; poets and astronomers alike must work with similar raw material. Imagination aids the poet in conjuring up alien vistas, while the astronomer is assisted by sophisticated devices that split the light into a thousand frequencies, unravelling the most intimate nature of the stellar matter. It is a rare astronomer who is capable of clearly communicating his esoteric discoveries, his knowledge of what really constitutes a star, to the uninitiated audience; and, more, of enthralling his readers, who are too frequently abused by sycophantic disciples of pseudo-science.

In *One Hundred Billion Stars*, Mario Rigutti has escaped from the dreary level of popular science writing that permeates the pages of numerous magazines. He does not give us breathless prose, studded with superlatives, yet manages to transmit a genuine awe for the starry heavens. Professor Rigutti's style is chatty and

Advertisement

informal. He writes as though he were telling a bedtime story to a child, and succeeds in captivating people of all ages. He takes occasional potshots outside the world of astronomy, rightly believing that practising scientists cannot ignore social and political issues. His unassuming approach and his evident sincerity provide a refreshing perspective on the world of the stars. Due credit should go to the translator, Mirella Giacconi, who appears to have ably succeeded in capturing the spirit of the original Italian text.

The book commences with a discussion of the Sun, a logical place to begin. It continues with the inner planets, and then

comes the major section, on our galaxy. Most of what is known about our galaxy is told, not too much progress having been made in stellar evolution since 1978 when the Italian edition appeared. The new discoveries have certainly not changed our overall view of our local celestial environment, and the book does not suffer too adversely from this shortcoming. The author rightly concentrates on capturing the essence of the underlying truth behind those glittering points of light in the sky, and, by and large, he succeeds. □

Joseph Silk is Professor of Astronomy at the University of California, Berkeley.

Pi in the sky

John D. Barrow

Constructing the Universe.

By David Layzer.

Scientific American Library: 1984.

Pp.313. \$21.95, £15.95. (Members of the Library only.)

Stephen Hawking's Universe.

By John Boslough.

William Morrow: 1984. Pp.158. \$12.95.

DAVID Layzer's addition to the Scientific American Library is extremely ambitious. It aims to tell the story of cosmological reasoning from the pre-Socratics to the present day. Not that it is just a history book: the author's approach is distinguished by his inclusion of accounts of the mathematical reasoning used by the past masters in their classic investigations. In this way one is able to see the actual logic of scientific development, a far better approach than the stream-lined résumé which uses modern concepts to tidy things up retrospectively. Layzer also enlivens the story by commenting on the discussion and dissent that surrounded the introduction of new concepts; for example, Berkeley's attitude to "fluxions". These early chapters dealing with historical developments are carefully written, beautifully illustrated and well-integrated into a discussion of the scientific method which runs through the text.

The next part of the book provides a well devised description of special relativity and a brief survey of the key concepts involved in the formulation, verification and application of general relativity. With all this groundwork laid out so thoroughly I expected to find as its climax a superb final portion of the book devoted to modern cosmology. I was disappointed.

The author makes much of the importance of correctly formulating the Cosmological Principle but then does so in a misleading and unappealing way by defining it to mean that "the universe itself has the same spatial symmetries as the laws that govern its structure and evolution". This statement is said to be "axiomatic",

which is unfortunate, since it is clearly false. Furthermore, the claim that any statistical formulation of the Cosmological Principle must be rejected because it implies preferred places is also illogical. We can only make scientific statements about the observable portion of the Universe and there may well exist "preferred" places where observers are most likely to be found. The "many-bubble" inflationary universe models considered recently need not satisfy any "axiomatic" cosmological principle.

However, my main worry about the book arises from the author's cursory, and to my mind, dishonest treatment of the observational evidence for the hot big bang model — a topic which ought to have been extensively and fairly discussed in a book such as this. The author displays "predictions" of the cosmic abundances of helium and deuterium from the hot big bang model supposedly made in 1965 and 1984, along with the "observations" of these abundances at those same dates. Virtually all the information provided is incorrect. The presently observed cosmic helium mass fraction does not lie in the range 10–21% as claimed, nor did it lie in the range 26–28% in 1965. As far as I am aware, there were no big bang deuterium predictions in 1965, nor were there any extraterrestrial measurements. Nor are the predictions any different, in principle, over this time span; they are only made to look very different by the author's assumption that the present baryon density must be ten times higher than we believed in 1965 because of the flat rotation curves of spiral galaxies. But if the dark matter responsible for that state of affairs is non-baryonic, it does not deplete the deuterium abundance in the way Layzer claims. Further, quite a lot of significance is attributed to the Woody-Richards measurement of a distortion in the microwave background even though its statistical significance has since been downgraded.

Having presented this brief caricature of the hot big bang model, the author then devotes 22 pages to detailed exposition of his own speculative ideas developed in a student's thesis in 1973. These ideas regarding a non-primordial origin of the

ANSWERS

to your
questions about
nuclear energy



call
or
write
today
for a
free

catalog of
answers and facts
about nuclear.

AMERICAN NUCLEAR SOCIETY

555 N. Kensington Ave.

La Grange Park, IL 60525

(312) 352-6611

background radiation are interesting, but I think that they are given misleading and misplaced prominence in a book of this nature; after writing with unusual care and clarity for three-quarters of the way, it seems to me that Layzer has over-indulged himself. As a volume in the Scientific American Library series, a large number of people will read this book. It is a pity that having done so, they will know nothing of the exciting observations and theories that underpin our most recent ideas about the construction of the Universe. This is a book that could have been outstanding — what a sadly wasted educational opportunity.

It is difficult to know what to say about John Boslough's "biography" of Stephen Hawking, written, I gather, without its distinguished subject's blessing. Certainly, it is written in a style reminiscent of certain dubious tabloid newspapers. The book is built around a collection of quotes and anecdotes gleaned at first- and second-hand from sporadic conversations and interviews, but like many pieces of semi-sensational journalism the author is interested only in the extraordinary and the remarkable. This approach may work well when reporting the Cup Final but in any work of biography it is disastrous. Here, the real picture of Hawking's character and ideas is superseded by a patch-work of aphorisms and anecdotes that are atypical, for indeed that is why they are so memorable. Why has the author not interviewed Hawking's many collaborators and associates with the goal of presenting an accessible but thorough and accurate description of his way of thinking? Accessible need not mean trivial.

The author's handling of the scientific ideas involved is unreliable and I am sure, to anyone not completely familiar with what he is trying to explain, unintelligible. Buzz-words and acronyms spontaneously appear without explanation — "GUT's", "flatness", "inflation", "Guth's scenario" — while the melodramatic style leads to errors and absurdities: Einstein's special relativity paper was not a "bomb-shell"; he did not show that "time did not always flow from past to present"; the concept of a black hole horizon did not arise as a consequence of Hawking's Area theorem, nor is its surface area the only externally observable property of a black hole. Then there is the description of how Einstein discovered the mass-energy formula: "Einstein came to the conclusion that $m = E/c^2$. From there it was just a simple algebraic step to the most famous equation in history, $E = mc^2$ ".

Perhaps some will enjoy this book. I didn't. Stephen Hawking is one of the world's most remarkable scientists and I hope that a biography worthy of his achievements will one day be written. This, I am afraid, is not it. □

John D. Barrow is a Lecturer in Astronomy at the University of Sussex.

The best lab in the world?

Nevill Mott

Three Degrees Above Zero: Bell Labs in the Information Age.

By Jeremy Bernstein.

Charles Scribner's Sons: 1984. Pp.241. \$17.95.

BELL Laboratories are envied worldwide, and Bell scientists know it. Of what other industrial laboratory could the vice-president for research write:

... if I hire the fifth-best theoretical physicist in the country I can get away with it, but if I hire the fifteenth-best I am wasting my time. We are looking for number two or three, and we want to compete with the best universities for number one.

One can sometimes notice at conferences that the Bell scientists sit and talk together, and those from other laboratories sit elsewhere. What other industrial laboratory has had seven Nobel Laureates since the War? And certainly no other laboratory has been funded by the company (AT&T) with a monopoly of telecommunications in the United States.

Jeremy Bernstein's book gives a picture of Bell Labs by describing certain projects and some men and one woman at the centre of its achievements. It is written, too, at a moment when the future is in doubt, and uneasiness about the coming years colours many of the articles in it. Just two years ago an agreement was signed between the Justice Department and AT&T which split up the telephone company but allowed it to retain Bell Labs. Fears are expressed that fundamental research may be squeezed, results expected too quickly, supervision become too close and that Bell Labs will become "like any other industrial laboratory". Some of the scientists who feature in the book are optimistic but the anxiety comes through only too clearly. Those who have watched events from across the Atlantic have wondered whether it was worth while, so as to introduce more competition, to put at risk the quality of, perhaps, the best laboratory in the world.

I do not know whether Bernstein wrote this book because of what had happened to the telephone company; he does not say so, and it is not his main theme, though perhaps the most pertinent one. His aim is to tell a story. He writes of the early days of the laboratory, and then, inevitably, about the transistor.

This story has been told elsewhere, for instance in Braun and MacDonald's *Revolution in Miniature*, published by Cambridge University Press in 1978, and the author acknowledges what a fine study that book is. Reading Bernstein's own account, it seems certain that no organization but Bell could have made the breakthrough, at any rate not until years

later. Even before the War, the need to replace the vacuum tube was recognized. The achievement, through the understanding of theory coupled with the most practical advances in materials science, could have occurred only in a place where the need was urgent and where all these skills existed. Even so, it was slow to come to fruition. The discovery was in 1947; it was the end of the 1950s before the electronic revolution was really underway.

Perhaps the most interesting pieces in the book are the vignettes of some of the people active now. The book had its origin in a discussion between the author and Philip W. Anderson, who shared the Nobel prize in 1977 with his former teacher J.H. Van Vleck and the present reviewer. An outstanding paper of Anderson's, which started the applications of solid state theory to electronics in glasses, was called "Absence of Diffusion in Certain Random Lattices", published in the *Physical Review* in 1958, and is referred to in the book's "Selected Bibliography". As the book is meant for non-specialists and indeed for readers who are not scientists at all, I should warn them that even for experts this is an exceptionally difficult paper; Anderson himself said it was "often quoted but never read". Its results were frequently disbelieved, but once accepted it was extraordinarily illuminating and, with a time delay of nearly ten years, sparked off a burst of theoretical work on non-crystalline semiconductors. Of course this work continues with great vigour at present.

Of interest too is the sketch of Anderson's career. It is surprising how hard it was for a theorist of genius, graduating from Harvard, to find a job in the immediate post-War years. But it was to Bell Labs he wanted to go, and he succeeded. Now, at sixty, he is something of an *éminence grise*, with a chair at Princeton and a role as adviser to the directorate of Bell. He has immense influence in American physics.

The woman in the story is Suzanne Nagel, described in 1970 when she joined the laboratory as "a very attractive, 27-year old, auburn haired Ph.D". Now she is one of many concerned with the glass fibre business. Of her current work, she told the author:

I have been especially concerned with wringing the last drops of water from the fibers. Very small amounts of water — parts per billion — can cause attenuation of light in the spectral region where the lasers are emitting... Recently, by the way, we transmitted light for a hundred kilometers without the need for reamplification... The first transatlantic cable is due in 1988.

The title of the book refers to the temperature of the radiation background in outer space — the remains, so it is believed, of the energy dissipated in the Big Bang with which the Universe began 12×10^9 to 15×10^9 years ago. In fact the discovery of

this radiation is perhaps the strongest evidence that the Universe did begin in this way, rather than as in Fred Hoyle's theory of continuous creation. The reader who doesn't know Bell Labs may well wonder what this has to do with an industrial laboratory, but after the War it was natural for a communications laboratory to hire a radioastronomer, and the consequence was the discovery of the background radiation by Arno Penzias and Robert Wilson. The management allows that sort of thing, and was rewarded when these two won the Nobel prize, shared with Peter Kapitza in 1978. As I remember from 1977, a Bell Nobel prize can be used as a song of triumph.

The last chapter is about Penzias himself. His father, born and resident in Germany, was a Polish Jew. In 1938 the Germans started deporting to Poland those who did not have German passports. Penzias and his parents got to the Polish frontier, but the Poles would not take any more people. After a brief period of imprisonment the family was sent back to Munich and told to get out in six months or else go to Dachau. Fortunately they succeeded in emigrating, and came to America. Education in New York's elementary schools for a boy who started with no English was tough, but Brooklyn Technical High School, New York City College and Columbia University under C.H. Townes led to Bell Labs, radioastronomy and the Nobel prize. Penzias is now vice-president in charge of research, an advance which he regards as "nothing short of miraculous".

This chapter on Penzias tells most about how Bell Labs is run and about fears for the future. On management he says:

We have to understand, and I think our owners do, that it can take years and years before someone may be able to make something useful out of a piece of pure research. That we understand this is one of the unique strengths of this place. The few competitors we have, who hire people as good as we do, do not do as well as we do because they *manage* them too closely.

The italics are mine, and it expresses the secret of Bell Labs' success. May it continue! Penzias's last words are: "I am a high pressure guy, and I didn't take this job to conduct a going-out-of-business sale". □

Sir Nevill Mott is Emeritus Professor of Physics in the University of Cambridge. In 1977 he shared the Nobel Prize for Physics with J.H. Van Vleck and Philip W. Anderson, whose work is described in this book.

Optics applied

Springer-Verlag have issued a second edition of Matt Young's *Optics and Lasers: Including Fibers and Integrated Optics*, a revision of the original work of 1977.

The book is an introduction to the engineering and applied aspects of optics, and has been updated to take account of advances in optical communications and the development of integrated optics. Price in paperback is DM 64, \$22.

Life, the Universe . . . everything

C.J.S. Clarke

The 4th Dimension: Toward A Geometry of Higher Reality.

By Rudy Rucker.

Houghton Mifflin: 1984. Pp.228. \$17.95.

To be published in Britain in April 1985 by Hutchinson.

CAN the picture of the world developed by modern physics really affect the way people act and think? Clearly Rudy Rucker believes that it can, and to this end he sets out to teach the reader to think four-dimensionally. He achieves his aims mainly through a careful dimensional expansion of E. A. Abbott's romance about the two-dimensional inhabitants of Flatland, published one hundred years ago, which he vivaciously continues in sections on "The Further Adventures of A Square". Initially Rucker develops the idea of our space (in the common, three-dimensional sense) being embedded in a four-dimensional hyperspace, which allows for the possibility of space being curved, the whole exercise being carried out in a light-hearted style with copious cartoons. When he passes to space-time he can then build on the reader's understanding of four-dimensional geometry to describe the relativistic picture of the Universe as a four-dimensional manifold. In addition he can combine the pictures of space-plus-time and of space-in-hyperspace to move on to the idea of space-time as a whole being embedded in a five-dimensional hyperspace. Finally, with the inclusion of ideas derived from quantum mechanics, he presents a picture of the world as infinite-dimensional, composed of the coordination of every inhabitant's separate perceptions of reality.

One great achievement of the book is that it should help to overcome the feeling of bemused awe by which non-mathematicians are often overwhelmed when multi-dimensional space is mentioned. On the one hand, the author succeeds in explaining clearly the way in which it can be useful to expand the mathematical description of space in terms of three parameters when further degrees of freedom are involved; on the other, he shows by discussing hyper-cubes and hyperspheres how the additional parameters can be built into a complete geometry.

But to convert the reader's heart and mind an author must do more: he must show that the higher spaces being built up are genuine geometrical wholes and not just arbitrary mathematical toys. In the case of special relativity Rucker conveys this well, showing that in the space-time of fast-moving particles the decomposition into space and time depends on what particle is used as the basis for a reference

frame and so is not intrinsic to the space-time itself: it is a true unity in which space and time have no absolute individual existence.

Though he is convincing on relativity, what are we to make of Rucker's higher space where the parameters being combined are as different in kind as (to take one of his examples) the raw-cooked and the sweet-salty axes for food? Surely here, whenever it is possible to define points in the space, it is also possible to decompose the space uniquely into the direct sum of the two constituent spaces. Before one can truly speak of a many-dimensional geometry, as opposed to the mere juxtaposition of lower-dimensional geometries, there must be some internal isotropy or symmetry between the dimensions.

This need for a symmetry between the dimensions becomes a crucial problem when it comes to the idea that our space-time is embedded in a hyperspace. One of the main growth areas in theoretical physics is the study of Kaluza-Klein theories where just this happens. But the symmetry relating spatial directions with directions into the hyperspace is a super-symmetry that does not preserve the normal fabric of matter, and the size of the Universe in the other dimensions is generally taken to be minute. When Rucker talks of the world floating around in a hyperspace, attractive though the idea is, one must doubt whether there is any real connection between his speculations and those of the theoreticians. The weakness of Rucker's argument, at a macroscopic level, is that there is no parameter that could be used to define the additional dimension.

Another difference between the author's approach and that of modern theory concerns curved space. Rucker stresses the picture of the curvature of space-time being thought of as the curvature of a hypersurface in a hyperspace. By implication that hyperspace is flat (otherwise one would pass to a hyper-hyperspace to explain its curvature), whereas the hyperspace of modern theories remains curved. Sadly, no formal advantage has been derived from embedding space-time in some hyperspace, and it is not even known what is the minimum number of dimensions that such a hyperspace would have to possess in order to accommodate a general space-time: the best so far achieved is to show that it is less than 84 dimensions (see, for example, my paper in *Proc. R. Soc. Lond. A* 314, 417-428; 1970).

But I am in danger of imitating the reviewer in the *Times Literary Supplement* who criticized *Wind in the Willows* as making a "negligible contribution" to natural history. Rucker aims not to write a monograph, nor indeed to present merely an amusing curiosity, but to make his ideas fire the reader's imagination and so make a difference to the way he thinks. To return to the question I posed at the outset, I believe that there is an interaction between

scientific ideas and the ideas that form the basic motivations for our thoughts and actions, including religious ideas, but it is easy to make the connection seem more inevitable than it really is.

To take one of the most salient examples from the book: does the relativistic view of time as simply one dimension of a four-dimensional Universe have any bearing on our attitude to our own death? For many the fear of death is the fear of non-

True, it is a valuable insight that one's life has worth as a whole, and as part of a greater whole; that "I am, as it were, an eye that the cosmos uses to look at itself" (p. 147). This insight seems to fit naturally with relativity. But we should be conscious of the presupposed philosophical realism that seems to underlie this particular argument. Is the space-time, the para-compact Hausdorff manifold, that parametrizes the motion of fast particles

Atkins's exposition, have only a ground and an excited state — that is, OFF and ON, except that ON may be both organized and unorganized). The simplicity ensuing from contemplation of only two states is notable, extensions to greater complexity being straightforward. For me, such analysis resolved apparent conflicts between savants as to whether or not crystallization from supersaturated solution establishes a correspondence between entropy and randomness (you must specify your knowledge of the system); at less subtle levels, the book will lift the hex on entropy traditionally suffered by tyros.

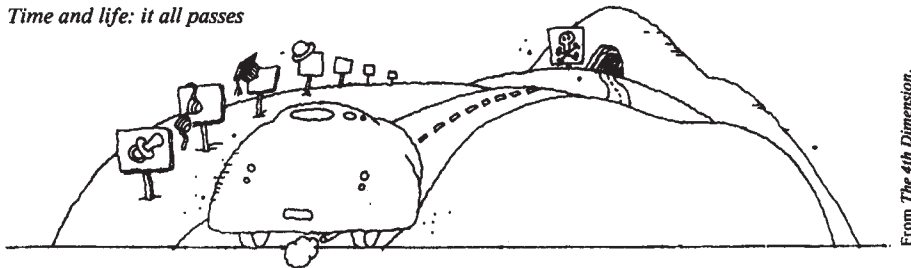
Here, engines of various nomination are unexceptionably cycled in commendably clear applications of the model of energy dispersion. Chemistry is shown to be controlled by free energy (though whether Gibbs's or Helmholtz's is relegated to the thermodynamics appendix). The emergence of complex biological systems (Life itself?) as a result of greater chaos elsewhere, is cogently argued. The so-called "Life Game" is outlined and there are several computer programs appended to play this and other exercises masquerading as games, a nice didactic ploy (also available, note, on discs).

Purists, however, will find much to balk at in the philosophical implications thus laid before them. For example the validity of contemplating the entropy of the Universe is severely questioned in current correspondence in *Chemistry in Britain*, and indeed has been over the years by many cognoscenti. Atkins's equating of entropy with randomness or chaos is justified by his meticulously numerical definition of randomness, but his use, as synonyms for energy dissipation, of such emotive terms as corruption, chaos and decay — perhaps in a striving to soften the austerity of traditional scientific prose — is questionable. And introducing "Jack" and "Jill" as engine operators mistakes the mental age of the reader, who must distinguish Atkins's "universe" (a large enough region encompassing observable change) from his "Universe", the Whole Thing; unluckily the illustrator has got the capitals wrong in the accompanying diagram (and elsewhere has colour-coded the fluctuations diagram incorrectly).

But this is still a lovely book, beautifully illustrated and presented, and clearly commensurate with its companion volumes in the Scientific American Library. It should engender affection as a sophisticated latter-day Hogben, complete with the social allusions implicit in the call to conserve, not energy, but entropy, and to use the virtuous entropy-conserving heat pump rather than crude and profligate combustion. If, for the purist, the net is cast too wide, it may nevertheless capture minds otherwise intent on escaping the chill waters of pure thermodynamics. □

David R. Rosseinsky is Reader in Physical Chemistry at the University of Exeter.

Time and life: it all passes



existence, as expressed by Unamuno when he wrote:

For myself I can say that as a youth and even as a child I remained unmoved when shown the most moving pictures of hell, for even then nothing appeared quite so horrible to me as nothingness ["Del Sentimiento Tragico de la Vida", quoted by R. Sorabji in *Time, Creation and the Continuum* (Duckworth, 1983)].

As argued by Sorabji in that book, this fear depends on a particular view of time in which there is some absolute distinction between past and future; for we were just as non-existent before we were born as we shall be when we are dead, and yet there are a great many who are indifferent to or intrigued by the former but appalled by the latter. In relativity, according to the most usual metaphysical gloss on the subject, not only is there symmetry between the future and the past of any event (so that a horror evoked exclusively by future non-existence is shown to be irrational) but it is not even possible to say that "at time t I was/will be non-existent", because there is no absolute meaning to "time t ": the statement that I shall not exist (anywhere) in the year 2060 is metaphysically equivalent to the statement that I did not/shall not exist (anytime) in Mongolia, except for the first being more certain to be true. Every part of the space-time universe is equally tenselessly existent.

Rudy Rucker writes in this book that

Instead of thinking of myself as a decaying bag of meat, I can think of myself as a part of eternal spacetime. This is a way to cheat death. Instead of identifying myself with my specific body pattern, I identify myself with the block universe as a whole [p. 147]

and because of the impossibility of defining a "now", a "time t ", in which to localize existence

The idea of the block universe is, thus, more than an attractive metaphysical theory. It is a well-established scientific fact... Spacetime is a single unified whole, and the passage of time is just an illusion [pp. 149, 155].

actually the same thing as the space and time in which we develop our lives? Or is the manifold more "real" than the time of our experience just because it relates best to physics? I plead guilty to being a confirmed realist myself; but the times when I have been most aware of reality bursting through delusion have been in human, rather than mathematical, experience. □

C. J. S. Clarke is a Lecturer in the Department of Mathematics at the University of York.

Law, and prophecy

David R. Rosseinsky

The Second Law. By P. W. Atkins. *Scientific American Library*: 1984. Pp. 230. \$21.95, £16.50. (Members of the Library only.)

A PRECEPT for both admirers and critics of Atkins's book could well be "Thou shalt not make unto thee any graven image". *The Second Law* encompasses, in essentially non-mathematical form, the rigorous to the reflective, but Atkins firmly imputes mechanism to every process. C. P. Snow's contemporary Renaissance Man would know both Shakespeare and the Second Law, and even a popular entertainment by Flanders and Swan some time ago properly enunciated the Lex Secunda, before mangling it in the chant, "Heat is work and work is heat". Correctly, work may all be simply converted into heat, but heat cannot all be simply converted into work. The complexity of fuel plus machine necessary to have work done contrasts with the ready processes available for squandering energy as mere heat.

In explanations of great clarity, Atkins relates spontaneous macro processes to particulate systems where energy becomes dispersed over localized sites (which, in

Nuclear science in war and peace

Philip Davenport

The Making of the Atomic Age.

By Alwyn McKay.

Oxford University Press: 1984. Pp.153.

Hbk £12.50; pbk £3.95.

FORTY years is an expedient interval after which to commemorate an important international event. The main figures, heroes and villains alike, tend to have expired, while there are enough younger participants still around to ensure authenticity and fair play. This year we have remembered the invasion of Europe on D-day 1944; next year will see the fortieth anniversary of not only the end of the war in Europe but also the surrender of Japan. To the vast majority this last event came as a complete surprise and with devastating suddenness. We may expect next year a resurgence of public interest in the discoveries and developments which led to the production of the first nuclear weapons and a corresponding demand for non-specialist books on the subject.

Dr McKay's book is well suited to meet this need, especially in its reasonably priced paperback form. He tells the story of early nuclear science from the discovery of radioactivity to the concept of nuclear fission with little technicality and much human interest. As befits a chemist, he stresses the importance of chemical expertise in these pioneering researches,

when minute quantities of nuclides had to be isolated and identified. As the reality of nuclear fission became substantiated the behaviour of scientists diverged, some rushing to establish priority of publication, others covertly seeking to corner supplies of uranium ore. Then followed caution and voluntary censorship as the possibility of chain reactions and hence nuclear explosives became apparent.

The middle chapters provide a readable and fast-moving description of the Allied wartime activities and the Manhattan Project which produced the bombs dropped on Japan in 1945. For those who have read the official histories and the biographical accounts of this enterprise there will be few surprises. The history of the German efforts, hampered by Nazi politics and anti-semitism, is also quite familiar. But the general public must be largely unaware of the wartime existence of a Japanese atom-bomb project which they managed to conceal so successfully that it did not come to light until the mid 1970s.

In his final chapter Dr McKay surveys world energy resources, including the so-called "alternatives" of solar, geothermal and tidal energy. He suggests that Divine providence has enabled man to extract energy from uranium, so that we can survive when fossil fuels become exhausted. I find it a pity that he expresses no such confidence in the future role of nuclear fusion. □

Until his retirement Philip Davenport was engaged in research at Culham Laboratory and is now its honorary consultant on historical matters.

Physics and faith

Ziauddin Sardar

Ideals and Realities: Selected Essays of Abdus Salam.

Edited by Z. Hassan and C.H. Lai.

World Scientific/Wiley: 1984. Pp.369.

Hbk £32.20, \$37; pbk £16.10, \$18.50.

A PHOTOGRAPH in *Ideals and Realities* shows Abdus Salam, wearing a "fez" cap and academic gown, looking rather reflective at receiving his honorary degree of Doctor of Science at the Aligarh Muslim University, Aligarh, India. The Aligarh Muslim University was founded, exactly a hundred years ago, by Sir Sayyid Ahmad Khan, a man who not only looked a bit like Salam but also shared his passion for learning. Sir Sayyid, a deeply religious man, was an advocate of "modernism" in the Muslim world. Abdus Salam, also highly religious, has pioneered science in Muslim countries. Both men have built institutions and have been honoured internationally for their endeavours. Both have become controversial figures within their own communities — Sir Sayyid for his faith in

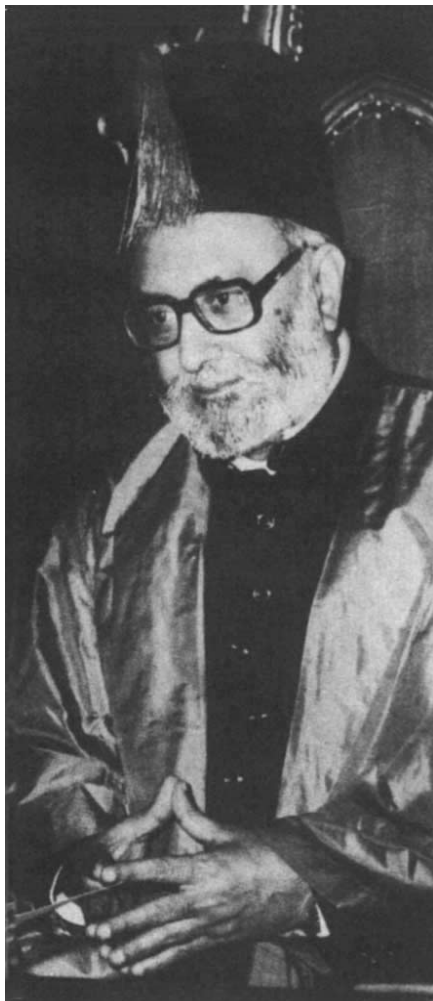
Westernization and Salam for his heterodox religious beliefs.

But there is a radical difference between the two men. Sir Sayyid was a born optimist: although he was upset about conditions in Muslim societies, he believed in the future. Salam is a pessimist: he plods along despite his increasing conviction that science will never take root, given present trends, in Muslim societies.

It is his pessimism that comes out most strongly in this collection of essays. Written over the past 25 years, they cover wide ground, ranging from international cooperation in science, science in the Muslim world and at the International Centre for Theoretical Physics (his own institution), to "perspectives" on physics. The first section of the book contains four introductions to "Salam the Man".

Although there is a great deal of repetition of material, one is quite happy to bear with this because of the conviction and insight that each essay carries. It is fascinating to read the book from cover to cover and see the shaping of Salam's outlook. The scene is set by the brief opening essay entitled, "The Less-developed Countries: How Can We Be Optimists?". Written some 20 years ago for Nigel

Calder's anthology *The World in 1984*, the article gives the reasons why Salam is so "utterly pessimistic": "the agricultural production of all but the richest countries is static"; "there are none among the rich nations willing enough to sponsor a fair price structure for the commodity market"; "the battle to keep the trickle of foreign aid programmes flowing becomes



Abdus Salam — how can we be optimists?

fiercer and fiercer"; men of "passionate fury" are not coming forward from the Third World to fight poverty; and where there is a realization of the struggles ahead, "it has not been purposeful enough yet to bring down the internal, social and the organisational barriers to be able to defy external pressures". From here on, in essay after essay, powerful but gloomy, he pleads the case for the poor — arguing for more spending on science and help for scientists in developing countries, making suggestions that this or that institute should be set up to promote science in the Third World, that new technological universities should be established and that the policy makers of developing countries should pay more attention to pure sciences and the development of indigenous talents and resources.

Yet, in his own life Salam has nothing to be pessimistic about. The young lad from a small Pakistani village, Jahlam — not

renowned for producing great scholars — has reached the heights of scientific achievement. His life is full of lucky turns. He could have easily ended up as a scholar of Urdu or Persian; or worse still, grazing in the offices of the Indian civil service, on which his father was very keen. On his return from Cambridge, after finishing his doctorate, he could well have disappeared into the anonymity of the colonially inherited educational institutions of Pakistan. But fortune was on his side and he escaped.

Perhaps fortune has little to do with it. It is Salam's hard work and remarkable intelligence that eventually brought him the Nobel Prize (awarded in 1979 for his work on the theory of electro-weak force). Or perhaps it is his sincerity of purpose and deep humility combined with a love for beauty and learning that is the driving force behind his achievements. Whatever the reason, Salam could not have overcome his own lugubrious nature without deep faith. Over the years, as this book shows so convincingly, his faith in science as the ultimate pursuit of objective truth, and his belief in Islam as the world-view of compassion, have increased not diminished.

Salam has achieved a remarkable, personal synthesis of Western science and Islam. It manifests itself in all that he preaches and in his own work. His involvement with symmetries in physics, he has said, stems from

my Islamic heritage for that is the way we consider the universe created by God, with ideas of beauty and symmetry and harmony, with regularity and without chaos. The Koran places a lot of emphasis on natural law. Thus Islam plays a large role in my view of science; we are trying to discover what the Lord thought . . .

The fact that Salam was looking for unity in seemingly disparate forces of nature is considered by him to be part of his faith both as a physicist and a Muslim. In more than one essay, he quotes from the verse of the Qur'an that reads:

Thou seest not, in the creation of the All-Merciful, any imperfection
Return thy gaze, seest thou any flow
Then return thy gaze, again and again
Thy gaze comes back to thee dazzled, awed.

This is the faith of all physicists, Salam tells us. The deeper we seek the more is our wonder excited, the more is the dazzlement of our gaze.

Ideals and Realities provides a revealing insight into the mind of Abdus Salam. There is a great deal here that one can disagree with, argue against, even dismiss as too simplistic. But one cannot help being moved by Salam's compassion and conviction, his strong faith in science and in his religion, his concern for the developing countries, and by the facility with which he hatches one bright idea after another. My only hope is that his melancholy outlook is not contagious. □

Ziauddin Sardar is Director of the Center for Policy and Future Studies at East-West University, Chicago.

Keeping the gates

Philip H. Abelson

How to Edit a Scientific Journal.

By Claude T. Bishop.

ISI Press: 1984. Pp. 138. Hbk \$21.95; pbk \$14.95.

TO YOUNG scientists and others not well established, editors and their associated apparatus represent a fearsome and mysterious barrier to recognition and advancement. The continuing emphasis on publish or perish has enlarged the importance of the role of editors in decisions to accept or reject manuscripts. Hence, there should be a broad audience for a book that enables readers to look behind the scenes at the machinery that influences the fate of millions of scientists and the expenditure of billions of dollars. In addition, those scientists who are or may become part of the editorial machinery will find Dr Bishop's book helpful as a source of ideas and guidance about all phases of the editing of scientific journals.

The author brings to bear some 18 years of experience in editorial matters. For a number of years, he was editor of the *Canadian Journal of Chemistry*. Since then, he has had responsibility for 12 journals published by the National Research Council of Canada. These publications have circulations of 1,200 to 6,000, figures typical of most journals today. Dr Bishop's writing reflects his years of experience; that is, he seems conditioned to the problems faced by small journals with limited financial resources. Nevertheless, much of what he states is applicable to scientific journals of all kinds.

Discussion of the peer review system is especially good. In journals large and small the crucial factor in achieving good quality is the reviewing process. As Bishop points out, "All editors, and most authors, will affirm that there is hardly a paper published that has not been improved, often substantially, by the revisions suggested by referees". He further points out that,

One function of the refereeing system that is often overlooked is its indirect influence on the initial preparation of a paper. Established scientists write their papers with a critical sense that anticipates referees' questions. Without this subtle pressure in the background, there can be little doubt that the quality of presentation would deteriorate along with the content.

Because of the publish-or-perish syndrome, a minority of authors advocate abandoning the reviewing process altogether. They "regard all referees and editors as biased adversaries whose objectives are solely to reject, delay, or scoop all papers submitted to them". Bishop denies the validity of such views, stating that the common experience of editors is that "examples of intentional delay, biased reports, or unethical behaviour are extremely rare". It is, of course, one of the

important functions of editors to review the referees' reports to guard against delay or bias and to investigate allegations that material in a paper has been stolen.

The matter of ethics in the field of scientific publications has become of increasing concern. Furthermore, it is a matter on which editors have primary responsibility. Bishop devotes a chapter to the topic, pointing out that questions of ethics have been present throughout the history of scientific journals. However, scientists today are under enormous pressure to expand their bibliographies and a few cases of plagiarism and fraud have been uncovered. While calling attention to the situation, Bishop adopts a sensible view — on the matter of reporting fraudulent results, he notes that if the work is insignificant, the experiments may not be repeated nor will the paper necessarily mislead others; and if the claims in the paper are significant, the work will surely be repeated and its worthlessness exposed.

In the book the role of the editor is rather narrowly defined. Emphasis is on procedures to select and improve the manuscripts that are voluntarily submitted. But the active frontiers of science keep changing, and unless the content of a journal evolves it will become obsolescent and new publications will be created to fill the gap. It should be the responsibility of the editor to be alert to changes and if necessary to recruit manuscripts dealing with emerging areas related to the current content of the publication. The editor's flexibility is often limited by financial or political considerations. Thus if the editor is to be truly effective, he or she must be more than a custodian of manuscripts. Editors must be aware of — and participate in — decisions that affect their ability to produce a first-rate journal.

Although the basic processes of peer review and editorial decisions concerning manuscripts remain unchanged, the major journals in the United States have been modernizing their mechanism for accomplishing these procedures. They are incorporating the use of electronic devices — a matter that is barely touched on in this volume. Computer-assisted manuscript tracking and monitoring of reviewers' performance has been widely adopted, diminishing routine clerical tasks and improving selection of referees. The application of computer word-processing is beginning to transform the mode of transmittal of material from authors to editors. Already some manuscripts are being transmitted by diskette and ultimately the material will flow to the editor and to reviewers by computer networks.

However, though some of the mechanisms of editing may change, the basic relationships of authors and editors will remain largely the same. Bishop's book will continue to be a useful guide to authors and would-be editors. □

Philip H. Abelson is Editor of Science.

The sight of art

J.D. Mollon

Cameraworks.

By David Hockney.

Thames & Hudson/Knopf: 1984. Pp.280.
£30, \$50.

ADVANCES in visual science have more than once prompted a fashion in art. Pointillism and Cubism offer two examples. But whatever the theoretical commentaries provided by the artist, the connection between the perceptual theory and the art is usually loose or incomplete, and the art finally stands or fails on its own merits.

With *Cameraworks*, his new collection of photographic collages, David Hockney has joined the tradition of artists who offer a scientific rationale for their work. He has seized on an aspect of visual perception that has exercised experimental psychologists during the past two decades but yet is little understood.

When we look at a visually busy scene, such as Epsom Downs on Derby Day, the Grand Canyon from the North Rim or the interior of King's Chapel, it seems as if we have an immediate and richly detailed grasp of a panorama that subtends at least 100 degrees at our eye. Yet, if we fixate our eyes steadily on one point, we discover that our detailed perceptual grasp is limited to about 1.5 degrees, which corresponds to the tiny foveolar region of our retina.

We normally do not constrain our eyes to look fixedly at one spot, and, almost without the awareness of their owner, they dance around the scene, dwelling on any given point for about 250 milliseconds. The two eyes are yoked in their movements; and their angle of convergence (together with their focal length) is automatically adjusted as we look from background to foreground. From a train of successive fixations, our brain builds up for us an apparently detailed representation of an extended scene. If the scene is changing whilst we look, then the temporal variation is incorporated into our four-dimensional internal icon. And, to use the modern jargon, we seem to enjoy pan and scroll facilities in all the four dimensions.

Hockney is explicitly aware of this aspect of visual perception. He writes, for example: "There are a hundred separate looks across time from which I synthesize my living impression of you". He goes on to observe that it is not possible to look at a photograph, even an erotic one, for more than 30 seconds. And he argues, with a logic I cannot fully follow, that this is because a photograph, unlike a painting, takes very little time in the making. Hockney's solution is to take a multitude of snapshots of a scene or event and then to mount them in a collage.

He has foregone one major part of the aesthetic control open to the traditional photographer, the control of development

and printing. He either has used Polaroid snaps or else has sent rolls of exposed negative round to his local neighbourhood processors, "Benny's Speed Cleaning and One-hour Processing". (Benny's used to send standardized notices back with the prints, patiently explaining to Hockney what he had done wrong, "how I should try to center the camera on the subject, focus on the foreground, and so forth".)

Does it come off? I think one must consider separately those cubist-like collages where Hockney jumbles the component images and those montages where there is only minimal perturbation of the topography of a scene. In the former cases, where several snaps of a face or object, taken at successive moments, may abut and overlap each other, one's attention can be held for 30 seconds. But the reason, I believe, is the very opposite of the one advanced by Hockney. It is not that the collage expresses the normal experience of looking at a person or scene. Rather, the contradictions between local depth cues in one area and those in an adjacent area, the discontinuities of form, the variations of facial mood, combine to thwart and delay the perceptual mechanism that is struggling to construct a single, internally-consistent representation of the scene. Much as when we look at a still life by Braque, our visual system tries out successive hypotheses and is never able to attain coherence.

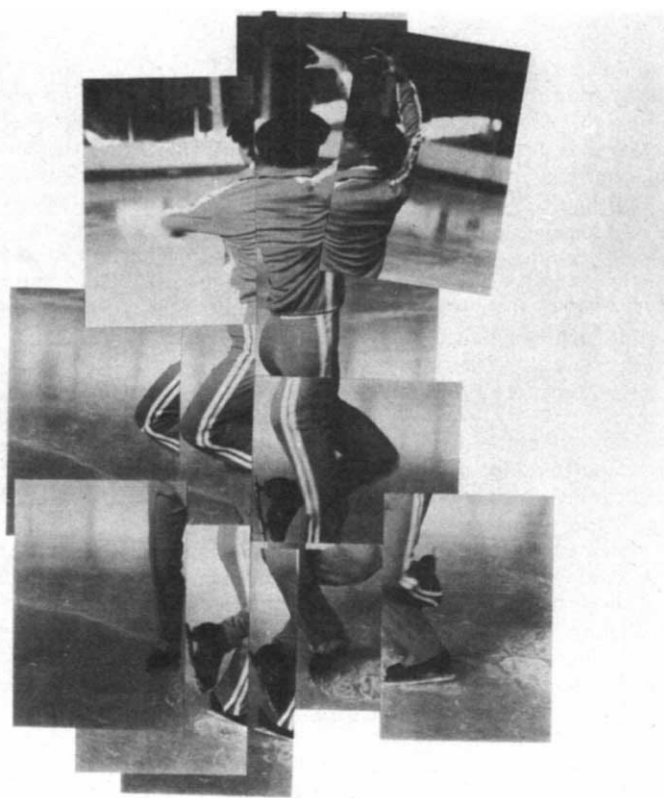
One of the best pieces in Hockney's collection — a portrait of the artist's mother at Bolton Abbey on a wet November day — is only spoilt by the intrusiveness of the collage. From the

glistening blades of grass in the foreground, past Mother in her blue-green oil-skin, to the grey-green graves beyond, the picture is marvellously coherent in its atmosphere; but the eye is interrupted by misfitting or missing bits of collage — and by the unevenness of Benny's processing.

One of the notable successes of the collage technique is in conveying the impression of rapid human movement. This can be well seen in the example reproduced here, where Hockney's spinning skater recalls classical Indian representations of the goddess Shiva.

In the case of the topographically coherent montages (for example his panoramic views of the Grand Canyon), I don't understand why Hockney doesn't save up the dollars he spends at Benny's and buy himself one of the specialist, slowly rotating, panoramic cameras that have been available since Friedrich von Martens introduced his *daguerreotype panoramique* in 1845. But a competent classical photographer, using a fixed lens, could often surpass Hockney's montages: there are single shots of the Grand Canyon by Andreas Feininger that capture more of the awesomeness and volume of the place than do Hockney's amateur montages of the same subject. And I could go on looking at, say, Feininger's "Ninth Avenue Elevated" (1940) for much more than 30 seconds — indeed for longer than it took me to flick through to the end of Hockney's *Cameraworks*. □

J.D. Mollon is a Lecturer in Experimental Psychology at the University of Cambridge.



The Skater. New York Dec 1982

Stars and sceptics

H.J. Eysenck

The Gemini Syndrome: A Scientific Evaluation of Astrology.

By R.B. Culver and P.A. Ianna.

Prometheus Books, 700 East Amherst St, Buffalo, NY 14215: 1984. Pp. 222. Hbk \$18.95; pbk \$11.95.

THE question of whether astrology is a science or a pseudo-science is not one which normally gives scientists a great deal of trouble. The well-publicized anti-astrology statement in the *Humanist* magazine of September 1975 was signed by 186 scientists, including 19 Nobel Laureates, and it probably mirrored the views of most people engaged in astronomy, physics, chemistry or psychology. Philosophers, too, have looked askance at astrology. Karl Popper, using his falsification criterion to separate out science from pseudo-science, grouped together astrology, psychoanalysis and Marxism as pseudo-sciences, arguing that none of them make falsifiable predictions.

Unfortunately there are difficulties with this position. The 186 signatories of the *Humanist* statement were fundamentally ignorant of the empirical evidence, and none have ever done any work in this field. Their declaration was clearly an act of prejudice; and while their verdict may be the correct one, the method used was not in line with the normal procedure of science. *Ex cathedra* pronouncements of this kind have no evidential value, and should be disregarded.

Karl Popper, equally, is not a good guide. I would agree that astrology, psychoanalysis and Marxism are pseudo-sciences, but not because they do not make falsifiable predictions. Every textbook on astrology is full of factual statements which can be tested, and indeed *The Gemini Syndrome* gives a good account of tests carried out to discover the truth or falsity of these predictions. For Popper the fact that most if not all of these predictions were in fact falsified may be irrelevant; for most scientists this will be more important in forming an opinion of the value of astrology than the simple truth that astrological predictions are falsifiable. Popper was equally wrong about psychoanalysis and Marxism; both make falsifiable predictions, and in both cases these predictions have usually been falsified!

Does all this mean that astrology is in fact a science? And can the same be said about psychoanalysis and Marxism? The answer surely must be that the criterion of falsifiability is a bad one for the purpose of demarcation. Much more important is the question of whether a given discipline has any positive contribution whatsoever to make, and on that basis astrology for the most part must be considered a pathetic failure. Culver and Ianna paint a dismal

picture of the inability of empirical research to verify any of the tenets of astrology, and in my book (with D. Nias) *Astrology — Science or Superstition?* (Maurice Temple Smith, 1982), I came to much the same conclusion. All the abracadabra of Sun signs, aspects, angles and whatnot cannot disguise the fact that horoscopes do not predict the future, do not give evidence about a person's character or intelligence, and cannot tell us anything about the diseases to which he is prone. Culver and Ianna write elegantly, with knowledge and fervour about these experiments, and it would be difficult to disagree with their wholesale dismissal of much of the evidence.

Unfortunately they throw out the baby with the bath water, as careless males are wont to do. In all the rubbish there appears to be a nugget of gold, namely the work of Michel and Françoise Gauquelin, and this is scandalously neglected by the authors. They do indeed give much space to the negative results reported by the Gauquelins, which are among the most impressive in the book, but their treatment of the positive results achieved is not only far too short to be intelligible, but is also inaccurate. In our book, Nias and I gave a whole chapter to the Gauquelins, and that is the least that their work demands; Culver and Ianna devote hardly two pages to it, and leave the reader completely unable to find out just what it was that they did, or how it should be evaluated.

Put briefly, the Gauquelins assembled, from library sources (biographies), lists of famous sportsmen, actors, scientists, medical men, military men, literary personalities and so on, and then looked at the planetary positions at the time of birth of these people. They found that certain positions (immediately after rise, and immediately after the upper culmination of a planet) were very significantly related to the various professions, Mars being important for sportsmen and soldiers, Jupiter for actors, Saturn for scientists and doctors, and so on. Note that the particular relationships between planet and profession are as predicted by astrological law, and that while the effects are small, the large numbers involved makes them very highly significant. Note further that the procedure was completely objective, with all the names of the individuals published in detail, so that all the results could be checked. Culver and Ianna are wrong in suggesting that the statistical methods used are questionable; this might have been true of the earliest attempts of the Gauquelins, but they have taken into account such criticisms, and their later work has been declared acceptable even by their severest astronomical critics. Moreover, there have been replications by others, a fact not recognized by Culver and Ianna, and these replications have been successful, such as that of the Belgian Committee headed by the Astronomer Royal of that country.

However, the Gauquelins' work has

New from
Churchill Livingstone

Marshall's Physiology of Reproduction

Volume 1: Reproductive Cycles of Vertebrates
Edited by G. E. Lamming

Fourth edition 832 pages illus
Cased £55.00

★ Detailed, up-to-date coverage of the reproductive cycles of: Fishes (*J. M. Dodd and J. P. Sumpter*), Amphibians (*B. Lofth*), Reptiles (*P. Licht*), Birds (*B. K. Follet*), Mammals; Monotremes (*M. Griffiths*), Mammals; Marsupials (*C. H. Tyndale-Biscoe*), Mammals; Non-primate Eutherians (*I. W. Rowlands and B. J. Weir*), Mammals; Non-human Primates (*H. G. Spies and S. C. Chappell*), Man (*S. S. C. Yen and A. Lein*).

★ Provides a framework for material to be published in the subsequent 5 volumes; each volume is to be edited by Professor Lamming in consultation with Sir Alan Parkes and an international panel of advisors.

Manual of Histological Techniques

John D. Bancroft and
Harry C. Cook

284 pages illus Paper £12.50

A manual of demonstration techniques, a practical companion to Theory and Practice of Histological Techniques.

Clinical Applications of Monoclonal Antibodies

Scientific Editor: E. S. Lennox

112 pages illus Paper £12.00

Describes the uses and potential of monoclonal antibodies including details of methods used to assess effectiveness and suitability. Fully illustrated including colour.

Also available as Volume 40 No 3 of the British Medical Bulletin.

Churchill Livingstone books are available from booksellers or, in case of difficulty, direct from the publisher at this address:

Sales Promotion Department,
Churchill Livingstone, Robert
Stevenson House, 1-3 Baxter's Place,
Leith Walk, Edinburgh EH1 3AF, UK.

Please send me:

- ☐ **Marshall's Physiology of Reproduction 4E Vol 1 Lamming** (443 01968 1) £55.00
☐ **Manual of Histological Techniques Bancroft & Cook** (443 02870 2) £12.50
☐ **Clinical Applications of Monoclonal Antibodies Lennox** (443 03042 1) £12.00

I enclose a cheque for £..... (made payable to LONGMAN GROUP LTD) or Please charge my Access/Eurocard/Barclaycard/Visa/American Express/Diners Club a/c

My credit card number is:

Signature

Name

Address

.....

Books may be returned within 28 days for a refund.

Churchill Livingstone 

44GM/NAT

Circle No.62 on Reader Service Card.

extended much beyond this. In collaboration with Dr S.B.G. Eysenck, they looked at the personality traits characteristic of all their thousands of subjects and found that these were related to the position of the planets in a meaningful and predictable fashion. They found very highly significant relations between the positions of the planets at the birth of a child, and those at the birth of his parents. They found that these relations obtain only when the birth was natural, but vanish when the birth was induced. This is not the place to give a complete description of all the Gauquelins' later work, but it very much reduces the value of *The Gemini Syndrome* that it fails abysmally to go into details about these experiments, that it does not mention more than a small part of them, and that it makes criticisms which are quite erroneous in the light of later developments. In view of the fact that the research of the Gauquelins is now recognized as the major positive support for astrology, the authors of a book such as this should have been especially careful to make sure to discuss it in detail; and if they were unwilling to agree with my own estimate of the value of these studies, they should have provided detailed criticisms which, if incorrect, could be refuted. Their failure to do so makes the book unacceptable as "a scientific evaluation of astrology", to quote the subtitle — a scientific evaluation does not leave out almost entirely evidence on one side of a question, while dwelling exclusively on that favouring the other side.

The work of the Gauquelins does not make astrology a science, but it does suggest that there are factual observations embedded in the mass of nonsense, and that it may be a legitimate task for science to dig them out and try to explain them. The extreme prejudice with which the Gauquelins' results were treated by many scientists does not constitute a good advertisement for the objectivity that scientists are supposed to manifest, and the way Culver and Ianna deal with the topic is unlikely to restore one's faith. □

H.J. Eysenck is Professor of Psychology Emeritus at the Institute of Psychiatry, University of London.

Behind the cue and the complex

Stuart Sutherland

The Hans Legacy: A Story of Science.

By Dodge Fernald.

Lawrence Erlbaum: 1984. Pp.241. Hbk \$29.95, £22.95; pbk \$19.95.

CLEVER Hans and Little Hans both achieved celebrity in the early years of this century. One was a horse whose feats included reading, arithmetic and telling the time. The other was a small boy who developed a fear of horses and who was analysed by Freud: his case plays an important role in the history of psychoanalysis because it was thought to provide some of the most striking evidence for the Oedipus complex. In *The Hans Legacy* Dodge Fernald retells the two stories on the pretext that they both illustrate the application of the scientific method.

Fernald's account of Clever Hans is of considerable interest, since the story has not been told in such detail in recent years. The wily horse's pretensions were exposed by Oskar Pfungst, a graduate student in psychology. Hans's owner, an honest man, had trained him to reply to questions by tapping out the correct number with his foot. In an investigation of great thoroughness, Pfungst showed that when the right number was reached people in the audience tended to raise their heads slightly and that Hans used this movement as a cue to stop tapping. His replies were random when he could not see anyone, while if Pfungst himself deliberately raised his head after the wrong number of taps, Hans made the corresponding mistake.

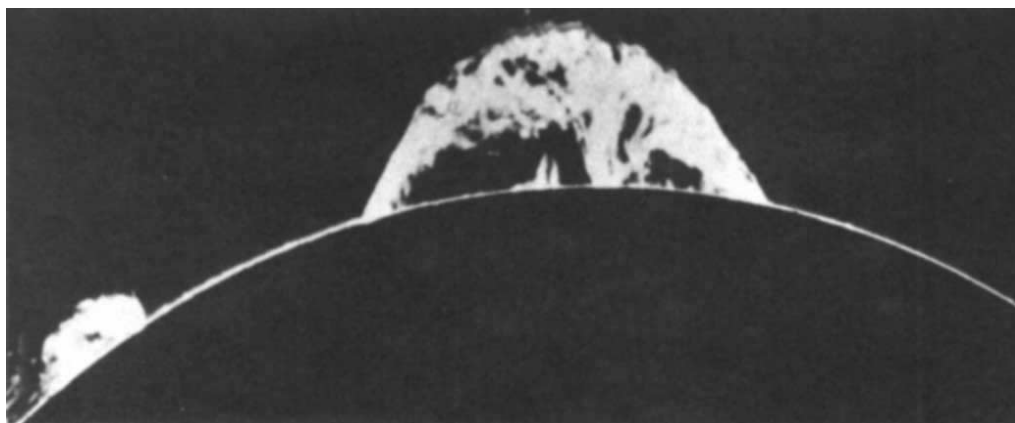
Much has been written about the other Hans, but regrettably Fernald does not cite one of the most important articles, that by Joseph Wolpe and Stanley Rachman who debunk Freud's claim that Little Hans's phobia provides evidence for the existence of the Oedipus complex. Little Hans himself claimed that his phobia stemmed from seeing a horse fall down in the street, a terrifying experience for a small boy. In

fact, Freud only saw his patient once during the analysis most of which was conducted at one remove by Hans's father, himself an ardent disciple of Freud's.

The father appears to have set out to convince the poor child that he had an Oedipus complex. Hans at first resisted suggestions that horses were a symbol for his father, and that he feared his father and wanted to supplant him in his relationship with his mother, but being a well-mannered and obedient lad he eventually gave a reluctant assent. Although Fernald brings out some of the ways in which Freud and the father misused the evidence, he does not reveal the extent to which they deceived themselves and insists on regarding the analysis as an example of the scientific method. If the cases of Clever Hans and Little Hans have anything in common, it is that both the horse and the boy were trained to respond in a certain way and that their trainers — Clever Hans's owner and Freud respectively — then proceeded to misinterpret the effects of the training.

Fernald frequently digresses. For example, he gives a brief but informative account of N-rays in whose existence French scientists believed for a time. The two main stories and the digressions are, however, more interesting than the lessons about scientific method that Fernald is determined to draw. He breaks up both investigations into the formation, testing and verification of hypotheses and he discusses the preparation of scientific reports, using as examples those written by Pfungst and Freud. Pfungst obligingly conformed to the standard format — "Introduction, methods, results, discussion"; Freud did not. Fernald's account of scientific method seems jejune, so much so that it is unclear at whom the book is aimed. No hints are given of the role of serendipity or hunch in science. He discusses scientific fame, but fails to acknowledge that it is often achieved not by those who discover important truths or who build interesting theories, but by those who make the most noise. □

Stuart Sutherland is Director of the Centre for Research on Perception and Cognition, University of Sussex.



Splendid outburst — the photograph, of a prominence on the limb of the Sun, is reproduced from *Secrets of the Sun* by Ronald Giovanelli.

The book is a highly illustrated account of the Sun and solar phenomena, and is published by Cambridge University Press (price £11.95, \$19.95).

Changing science for the time

Adrian Desmond

Georges Cuvier: Vocation, Science and Authority in Post-Revolutionary France.
By Dorinda Outram.

Manchester University Press: 1984. Pp. 299. £25, \$32.50.

Lamarck.

By L.J. Jordanova.

Oxford University Press: 1984. Pp. 118.
Hbk £7.95; pbk £1.95.

CUVIER'S aggrandisements have become a historiographic byword. The creator of stratigraphic palaeontology was also Napoleon's Inspector-General of Studies, Councillor of State from 1816 until his death in 1832, and Grand-Master of the Protestant Faculties. Darwinian propagandists derided him as a bureaucrat and bibliolater. Cuvier might have founded what Camille Limoges calls the "Macedonian Empire" of comparative anatomy, but he is still ritually censured in popular histories for leading the state opposition to Lamarck.

Dorinda Outram's *Georges Cuvier*, a deeply-reasoned attack on such rubber-stamp historiography, is a towering achievement in the "naturalization" of science history. Provocative, occasionally

emerges as a cosmopolitan liberal, scarcely religious, a friend of erstwhile Jacobins (who attended his wedding) and champion of working-class education — aspects at odds with the prevailing caricature. In a scintillating discussion of Cuvier's "catastrophist" *Discours* (1812), Outram asserts that — far from Scriptural kow-towing — Cuvier's abandonment of theological dogma led him to de-emphasize "causes" and endorse a historical descriptive approach. He demanded the "epistemological autonomy" of geology, staked its claim to the "new worlds of time" and, in an effort to conquer Lamarck's rival empire of conchology, founded the speciality of palaeontology, modelled after the new historical criticism. Outram's method is beguiling, and her imperial metaphors capture an expansionist era. This is a history of campaigns against competitors and annexations of nature: Cuvier's hostility towards Lamarck is a "territorial conflict" for control of the invertebrates. Patronage, property and power are the rewards for a successful play-off between social contingencies and natural resources.

For Outram, science is a "cultural commodity"; careers do not happen, they are made — hence her approach to Cuvier's rise, both in the *Muséum d'Histoire Naturelle* and *Conseil d'Etat*. She constantly questions the "heroic" view of affairs and shows that "disinterested" science is itself a myth of the scientific élite: a self-validating argument that requires a constant sanitization of history. That myth was even more essential in Cuvier's day. Given the uncertainties after the Terror (1793-1794), proclaiming a neutral ground could gain a *savant* immunity: "Whoever controlled the public exposition of such myths" of a politically-untainted science, she says, "was in a position to gain great public authority". The struggle to control the language of science climaxed during the 1830 Revolution, when Geoffroy challenged Cuvier's asocial ethos and used "populist rhetoric" to drum up radical support for his romantic morphology based on unity of plan.

With patronage networks ramifying through the political, banking and scientific worlds, one can see why Cuvier was forced to tailor his scientific claims to political expediences; society was volatile and his power-base unstable. He had to renegotiate claims with successive masters during the Revolution, Empire and Restoration, which entailed subtle modifications of his scientific stance. Outram reveals how, in turn, he met the challenge of revolutionary democratization, capitalized on the colonial predations of the revolutionary army, combatted Napoleonic "egoism", and retained power during the reaction, in part by modifying the rhetoric and content of his geology and zoology. In short, she has made splendid play of the mediating structures between the public language of science and

contemporary political events. Science, for all the propaganda, did not (indeed does not) inhabit a world of supralunary purity.

In *Lamarck*, L.J. Jordanova presents a corroborative view. Lamarck's debt to the anti-authoritarian *idéologues*, disliked by Napoleon, counted against him during the Empire, when the re-establishment of order became imperative. Jordanova has brought together the latest research on Lamarck, shaping it into the Oxford *Past Masters* format. She depicts the late

IMAGE
UNAVAILABLE FOR
COPYRIGHT
REASONS

Jean Baptiste Lamarck (1744-1829): legacy was not transformism, but taxonomy.

Enlightenment *idéologues* promoting a rational science of man and redirecting naturalists to the influence of environment. Lamarck's environmentalism, classificatory precision and attitude to transformed man are evidence of this debt.

But the enigma remains: Jordanova's Lamarck comes across as a minor nobleman in favour of the Revolution, a Heraclitean figure capable of majestic biological visions, who remained at heart a sea-shell systematist. His legacy to the following generation was not transformism, but taxonomy: even English evangelicals, who execrated his ungodly flights, paid homage to his conchological classification.

Lamarck is an admirable primer, and Jordanova's strength comes from her emphasis on the continuity between Lamarck's transformism and earlier theories of life and classification. Lamarck the *ancien régime* rationalist, old-fashioned chemist and obsessive meteorologist became Lamarck the transformist shortly after the Terror. He died a blind pauper in 1829, and was interred by Cuvier with a derogatory *éloge*. Cuvierian officialdom could certainly be heavy handed. But now Outram has taken us behind the politicking which made it so essential in a turbulent age. □

Georges Cuvier (1769-1832): play-off between science and political expediency.

conjectural, it is always absorbing. Outram has colonized terrain between cognitive sociology and psychology, dissecting the complex relationship between the individual, state and science that was forged during the French Revolution. This is not a study of Cuvier's comparative anatomy, but of his public persona as a *savant* — not the textbook history of transcendent genius, but a living biography of a man forming alliances and entering into "micropolitical" negotiations to advance his career.

The ossified image of Cuvier is that of the stern "legislator of science", arrogant, unyielding and conservative. Here he

Adrian Desmond is an Honorary Research Fellow in the Department of Zoology and Comparative Anatomy, University College London, and author of *Archetypes and Ancestors: Palaeontology in Victorian London 1850-1875* (Blond & Briggs, 1982/University of Chicago Press, 1984).



International code of zoological nomenclature, 3rd ed.

This important reference work sets out the internationally accepted rules for the scientific naming of animals, both fossil and living. Extensively revised, with enlarged Glossary and Index, it is required by all engaged in the execution and interpretation of taxonomic research. January 1985. 320pp (approx). Hardback, 0 85301 003 X. £15.00 Published by the International Trust for Zoological Nomenclature

Biology of Opisthobranch molluscs, Vol. 2.

T. E. Thompson & G. H. Brown

As in Vol. 1 every effort has been made to verify previous descriptions, and to compare British representatives with their European and North American equivalents. Dichotomous keys to species are provided and original colour plates from life depict most British species. 1984. 280pp. 41 plates (35 colour). Hardback, 0 903874 18 0. £39.00 Published by the Ray Society

Australian mealybugs.

D. J. Williams

This authoritative work is both an identification guide and review of current knowledge on the 196 species covered. It will aid correct classification and thus promote the use of appropriate biological controls where these are known. It will be particularly valuable to agricultural and biological control research workers. February 1985. 416pp (approx). 177 figs. Hardback, 0 565 00953 2. £40.00

A Revision of Dalziel's Useful Plants of West Tropical Africa, Vol. 1. H. M. Burkill

This work will eventually provide a definitive list of all West African Plants known to be used by man. Vol. 1 covers families A-D: each plant is fully described, its uses and location given, and bibliographical sources cited. An important reference source for all involved in agriculture, horticulture, medical science, food and plant sciences, etc. Early 1985. 900pp (approx). Hardback, boxed, 0 947643 01 X. £70.00 Published by the Royal Botanic gardens, Kew.

Peonies of Greece.

W. T. Stearn & P. H. Davis

A definitive and profusely illustrated monograph representing an important contribution to knowledge of the peonies of Greece, and the Mediterranean region as a whole. 1984. 136pp. 15 colour plates. 37 figs. Hardback, £28.00 Published by The Goulondris Natural History Museum, Greece.

The Catalogue of meteorites, 4th ed. A. L. Graham.

A. W. R. Bevan & R. Hutchison

This new edition covers all well-documented meteorites known to January 1984. Each entry includes a full description of the meteorite, details of its retrieval, and its classification where known. Especially useful is the list of prepared sections of meteorites available for loan in institutions other than the BM(NH). 1985. 400pp (approx). Hardback, 0 565 00941 9. £35.00 Copublished with the University of Arizona Press.

Supplement to Mammals of the Palaearctic Region. G. B. Corbet

This supplement brings together data published during the last decade which significantly affects the classification of mammals in the region and incorporates an extensive Bibliography. 1984. 46pp. Paperback, 0 565 00944 3. £6.00.

All orders to:
Publications Sales, British Museum (Natural History), Cromwell Road, London SW7 5BD.

Mendel in his place

Robert Olby

Mendel.

By Vítězslav Orel.

Oxford University Press: 1984. Pp. 107.
Hbk £7.95; pbk £1.95.

AS AN acknowledged world authority on Gregor Mendel (1822–1884), and an enthusiast for the history of genetics, Dr Orel was an appropriate choice of author for this 100-page biography in the *Past Masters* series. Only one full-length biography of Mendel has ever been written, that by Hugo Iltis, published 60 years ago. Although Iltis researched his subject thoroughly, brought to light important documents and presented a very readable account, the investigations of the past two decades have thrown fresh light on the subject, particularly upon the context of Mendel's education and scientific research. As director of the Museum Mendelianum in Brno in Moravia, Dr Orel is well placed to present an up-to-date view of his subject. Moravia now forms part of Czechoslovakia but in Mendel's day it was part of the Austro-Hungarian Empire administered from Vienna.

In such a brief sketch the author has summarized Mendel's experimental work, its reception by his contemporaries and his subsequent posthumous recognition in 1900, without becoming involved in current discussions of historical interpretation. For the most part, these topics are therefore treated in the conventional fashion: Mendel was the founder of the science of genetics, he accepted the Darwinian theory of evolution and foresaw the value of his work to that theory. That such a view is considered misleading in some quarters is not betrayed in Dr Orel's text. Instead it is the broad social and political context in which Mendel worked that is emphasized. Moravia had been influenced by the Enlightenment and the French Revolution, and education and scientific research were seen as essential to its economic and social progress. Thus the Agricultural Society in Brno was concerned with the improvement of the breeds of sheep following the example of Robert Bakewell, and of fruit following that of T. A. Knight, and meteorological data were collected to provide a basis for accurate weather prediction. In these activities the Augustinian Monastery played a prominent part.

When Mendel entered the Monastery in 1843 the Abbot was President of the Agricultural Society and an enthusiastic supporter of scientific research. Indeed, so preoccupied with responsibilities outside the Monastery did the Abbot appear to the Bishop of Brno, that the Bishop urged the termination of all teaching in the surrounding schools, retirement of the Abbot and even closure of the Monastery. For-

tunately the Abbot's appeal to Rome against the verdict was successful. In the accompanying memorandum the Order's special privileges and the achievements of its members in the field of education were stressed. Included was a mention of Mendel's devotion to science, especially physics, his teaching activity and his studies at the University of Vienna.

In the book we learn more about Mendel's performance as Abbot. This rather unhappy subject is treated frankly without recourse to special pleading in favour of Mendel. Mendel was elected Abbot in 1868, a year after constitutional monarchy had been established in the Austro-Hungarian Empire and equal rights accorded to Austrians and Hungarians, but these were not extended to ethnic minorities such as the Czechs. This caused bitterness and aligned the Czech leaders with the conservative party of landowners and churchmen, which had opposed the constitutional reforms promoted by the Liberal party. Consternation was caused when Mendel, as the Monastery's delegate, supported the Liberals. His action annoyed Czech nationalists and the other ecclesiastical representatives.

There followed the unfortunate wrangle between Mendel and the Government over the increased taxation of church properties beginning in 1874. As the only Abbot in the whole of the Austro-Hungarian Empire to persistently refuse to pay the tax, Mendel brought sequestration upon the Monastery and increasing isolation and estrangement from friends upon himself. Nine anxious years of writing protest letters to the authorities took him away from science and contributed to the decline of his health.

Whilst emphasizing the social and economic factors in Moravia predisposing Mendel to the study of hybridization, Dr Orel does not undervalue the importance of debates in nineteenth-century academic botany over the nature of fertilization and the constancy or inconstancy of species. Unfortunately he does not devote space to a more extended account of researches and discussions of the theory of hybridization in Mendel's time, instead sweeping his readers on to Morgan, Muller, Avery, Watson and Crick, and genetic engineering. Yet anyone familiar with the writings of Naegeli (Mendel's chief correspondent) and Wichura, who published his study of willow hybridization a year before Mendel on *Pisum*, will know that there was a very specific theoretical context to the subject of hybridization which is crucial to understanding Mendel's contribution and its subsequent fate. The result is that Dr Orel's Mendel is introduced very effectively as a product of progressive trends in nineteenth-century Moravia, but when the portrait is completed we are confronted not so convincingly with a precursor of modern genetics. □

Robert Olby is Reader in History of Science at the University of Leeds.

Southern isolation

Barry Cox

Discoverers of the Lost World.

By George Gaylord Simpson.

Yale University Press: 1984. Pp.222.

\$32.50, £23.50.

Ancient Australia, 3rd Edn.

Revised by Rudolf Oskar Brunnenschweiler.

Angus & Robertson: 1984. Pp.342. £20.

BOTH of these books concern studies of the history of southern continents that were isolated for much of the past 100 million years.

In *Discoverers of the Lost World*, G.G. Simpson recounts the history of the discovery and study of the unusual mammals that evolved in South America. These had a particular significance in the development of the theory of evolution. The great bones of the giant ground-sloth first convinced Cuvier of the reality of extinction, while, over a century later, Darwin's discovery of extinct fossil relatives of living South American mammals was one of the first observations that led him to consider the possibility of evolutionary descent involving gradual modification.

Though Simpson describes these early discoveries in South America, his book is mainly concerned with those who spent a significant part of their scientific energies in searching for and working on its fossil mammal faunas. After the Danes, Lund and Winge, the first South Americans to take an interest in the fossils were the Ameghino brothers of Argentina, born in the mid-nineteenth century. The studies themselves were carried out by the elder brother, Florentino, but the hard and lonely work of collecting in the great empty spaces of Patagonia was the responsibility

of Carlos Ameghino. Florentino is remembered both for his quaint habit of naming new genera by combining the first and family names of other scientists (*Henricosbornia*, *Thomashuxleya*), and also for his stubborn belief that all the mammals of the world (including man) were derived from the fossil marsupials of Patagonia.

Despite the early involvement of the Ameghinos, the collection of the mammal faunas of the continent was for long dominated by such North Americans as Hatcher, Scott, Riggs, Patterson and Stirling. Only in the past 40 years have South Americans, Paula Couto for example, started again to take over that role.

It is certainly useful to have this compilation of historical studies, for a knowledge of history not only adds to our understanding of the past but gives us an extra perspective on the present. At the same time, I feel that the degree of detail is sometimes obsessive. Is it really significant that Darwin noted the latitude of Puerto Deseado at 47°S when in reality it is closer to 48°S? And is it always necessary to note the trivial spelling mistakes in the writings of nineteenth-century scientists? Similarly, Simpson's knowledge of language and pronunciation leads him to an excessive amount of parenthetical information on this aspect. I would also have liked a map (or maps) showing the general locations of some, at least, of the many sites mentioned, and some of Scott's excellent restorations of South American mammals would have appropriately adorned the text. So while I'm glad that the book exists, I didn't particularly enjoy it.

Ancient Australia is the second revision by R.O. Brunnenschweiler of Laserson's original work of 1954. The original book was intended for laymen, and for amateur naturalists and geologists, but Brunn-

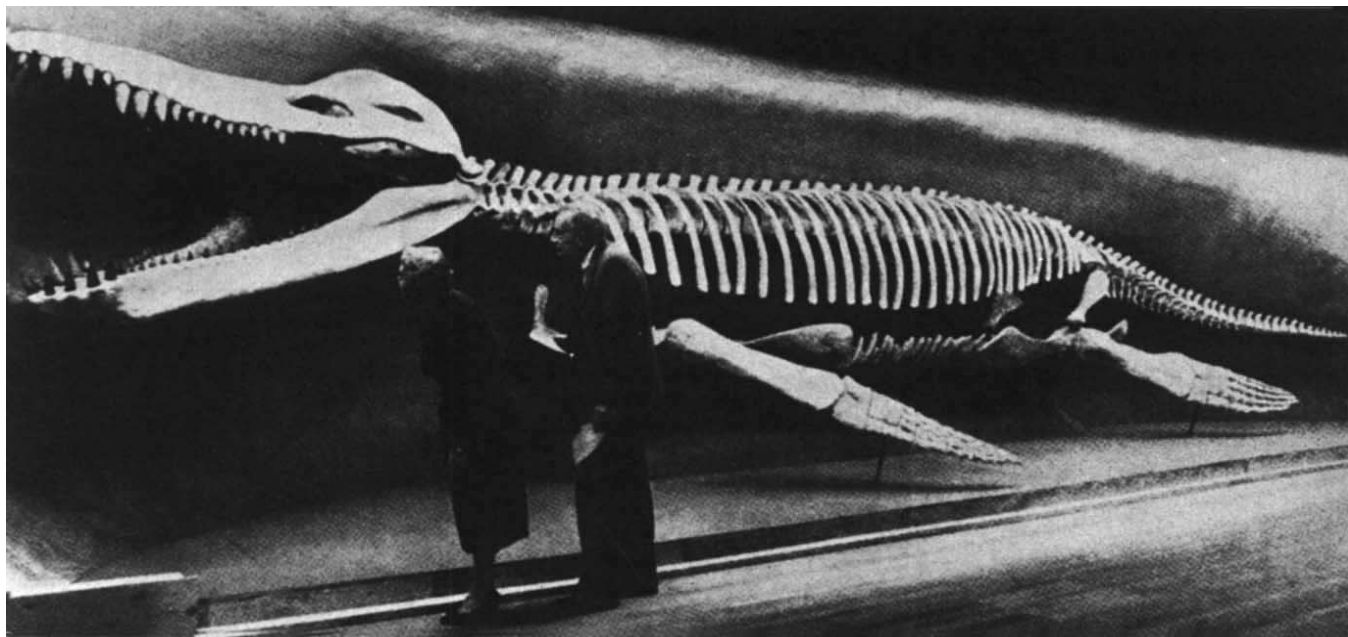
schweiler has broadened this aim to include professional scientists.

Apart from about 40 pages on the principles of geology, the nature of fossils and the variety of life, the book is made up of a series of chapters, each dealing with a single unit of geological time, from the Precambrian Eons to the Cenozoic Era. Each chapter commences with a map of Australia that shows the distribution of land, sea and lake, and then outlines the major basins of deposition, the various marine and non-marine formations, and the plant and animal fossils.

My main impression of the book is that it has been inadequately modernized. For example, the maps of Australia all show it as adjacent to the East Indies archipelago, and never as part of a Pangaea or Gondwanaland supercontinent. Similarly, the text refers to an Australian connection in the Triassic to New Caledonia and New Zealand "via a great arc of land around the Coral Sea", and states that the migration route for Triassic animals "must have been long and devious". Another example is the reference to the appearance of *Glossopteris* in "such widely separated lands" as Antarctica, South Africa, India and South America, and acceptance of the presence of *Glossopteris* in the northern continents.

The transformation from a stable-Earth to a plate-tectonic approach perhaps demanded a more radical re-working than the author was prepared to make. Nevertheless, its absence provides so old-fashioned a framework, lacking any of the fruitful integration of changing geography, latitude, climate, fauna and flora, that it greatly reduces the book's value to a palaeontologist or naturalist. □

Barry Cox is Head of the Department of Zoology at King's College, University of London.



From *Ancient Australia*.

Biting back — *Kronosaurus queenslandicus*, an extinct marine reptile 12.6 m long, found in Queensland and housed in the zoological museum at Harvard University. Examining the skeleton are A.S. Romer and Nelda Wright, Romer's research assistant.

Life among stones and bones

Jacquetta Hawkes

Disclosing the Past: An Autobiography.
By Mary Leakey.
Weidenfeld & Nicolson/Doubleday:
1984. Pp. 215. £12.95, \$15.95.

THE Leakey family, Louis, Mary and their sons Richard and Jonathan, have between them done more to reveal the early ancestry of the human family than the rest of the scientific world put together — or so it seems. Working together and separately in Kenya, Tanzania and Ethiopia, they have unearthed a prodigious number of primate fossils from Miocene apes to early forms of *Homo sapiens*. At the same time they, and Mary in particular, have made an important contribution to our knowledge of the earliest tool forms and their evolution.

I cannot resist thinking of the Leakeys as a unit, yet this book is very much a personal autobiography, opening with chapters on Mary's childhood with its constant travel dictated by her much-loved painter father. She had no regular schooling and drove away unhappy governesses. Can there be any other person today who has never passed a single examination and yet is heaped with honorary degrees and exalted academic honours of all kinds?

Through her father, Mary was already seriously interested in archaeology when, in 1933, she met Louis and they were soon irresistibly in love. Their association, and Louis's abandonment of his wife, rocked Cambridge and echoed round the archaeological world: Mary tells the story fully and with honesty. In the long run this scandal may have been a boon to science, for it meant that Louis, already known for his discoveries in East Africa, got no further support from his college and so was led to settle with Mary in his native Kenya.

In the late 1930s and through the war years they were ceaselessly active in exploration and digging, all the time so desperately short of money that their lives were hard and their work was handicapped. Elspeth Huxley observed "they live on the smell of an oil rag". Yet during the war their great discoveries began — outstanding among them Olorogesailie, where Acheulean handaxes lay massed on the ground like fallen apples, and the Miocene fossils of Rusinga Island in Lake Nyanza, where later Mary's extraordinarily sharp eye was to spot the first skull of *Proconsul*. Although this genus of ape has proved not to be a direct human ancestor, the find, made in 1948, was to lift Mary towards the heights of her independent career.

By this time Louis's staging of the First Pan-African Conference had brought African prehistory wide recognition and helped at last to attract sufficient funds for

the Leakeys to make regular excavations at Olduvai, the great gorge that slices through two million years of Stone Age history. After several seasons there, in 1959 Mary made the discovery which, she says, marked the second turning point in her life after that of *Proconsul*. This was the skull Louis was soon to announce in *Nature* as *Zinjanthropus*, a new genus — which he subsequently had to accept as a robust australopithecine.

"Zinj", alias "Nutcracker Man", brought the Leakeys popular renown and ample funds from the National Geographical Society. Ironically, Zinj had in fact been a disappointment to Louis, since he had always longed to find an early *Homo* which would expose the australopithecines as an abortive sideline of evolution. He was therefore overjoyed when, within a year and slightly lower in the same stratum, Jonathan came upon a larger-brained, more-gracile hominid. He was immediately recognized as having made the tools hitherto attributed to Zinj, and presently Louis dared to name him, again in this journal, as *Homo habilis*. Readers will recall the subsequent

controversy, which continues yet.

Cautiously, Mary recognizes *Homo habilis*, but it is just because she regards wrangles about our ancestral tree as premature — and odious — that she has preferred stones to bones, concentrating her research on material culture. She is still working, and the detection, not so long ago, of human footprints proving beyond doubt that hominids were walking upright over three million years ago, is one of her greatest triumphs.

Disclosing the Past is artlessly though agreeably written, and follows a chronological progress on the course of which science mingles with family and domestic affairs; in particular, the amazing variety of pets that have always shared Mary's homes here occupy many of her pages. I should judge that this very personal, informal chronicle, by a cool observer of both herself and events, will prove a valuable source for future historians of East African archaeology. □

Jacquetta Hawkes (Mrs J.B. Priestley) is an archaeologist and author of many books, among them *Mortimer Wheeler: Adventurer in Archaeology* (1982).

Truth on the ground

Derek Ager

Memoirs of an Unrepentant Field Geologist: A Candid Profile of Some Geologists and Their Science 1921-1981.
By F.J. Pettijohn.
University of Chicago Press: 1984. Pp. 260. \$25, £23.

ASA regular reviewer, I admit that I do not always read every word of all the books that are sent to me. Life is short and literature is long; I hope I do no injustice by my judicious skipping. But with this book I can honestly say that I read every word and they were all fascinating.

From the title I expected a stout defence of field geology by one of its greatest protagonists, and I presume I was asked to review the book because of remarks I published recently to the same effect. In fact I drafted a tentative opening to this piece saying that I felt like the madam of a brothel being asked if I approved of sex. That still applies and is exemplified by Professor Pettijohn's cry "Why, oh why, can't geology be taught where geology is and not in the lecture hall?". But although the title indicates the author's basic philosophy, the book is much more than that.

Francis Pettijohn was in at the birth of most of the great developments in geology this century, from geochemistry and isotopes to marine geology and basin analysis. In sedimentology he was truly the midwife and through the book we see the vigorous growth of that infant with its crucial importance to the rapidly expanding oil industry in the age of the motor-car.

Of course those of us like Pettijohn, who resist the urge to leap into black boxes, are likely to be labelled as old-fashioned and reactionary. When the thin section was introduced to geology by that great man H.C. Sorby (one of the few British geologists mentioned here), there was a similar resistance to "looking at mountains under a microscope". But fortunately geologists continued to look at mountains just as they now check their satellite pictures with what is quaintly known as "ground truth". What Pettijohn is saying is that there is no other kind of truth. At the same time, like the first-rate geologist he is, Pettijohn makes use of all the tools that are available to him. Thus he was a pioneer in the use of the aerial magnetometer for mapping the Lake Superior ironstones (it had only previously been used for detecting submarines).

But we start from the world of a small Mid-Western town at the beginning of the century, with an unpaved main street lined with hitching posts; the heroic age of American exploration had just ended, with the last of the covered waggons and the great gold rushes. Pettijohn's parents were school teachers in Wisconsin. His father later moved into local university administration and then to the University of Minnesota where Pettijohn began his formal geological education. His distinguished career unfolded through a series of universities, from Minnesota to Berkeley, to Chicago and to Johns Hopkins (the last move largely because Chicago had incredibly abandoned field geology). Like many academic geologists in the United States, he also worked for the oil industry and for the US Geological Survey.

There is in the book more about geolo-

gists than geology. The author seems to have known all the leading participants in the development of American geology during the twentieth century. The old departmental battles are fought over again and we are impressed or bullied by the Great Men; I was particularly taken by the portraits of A.C. Lawson, the "King" at Berkeley (who fathered a son at 88) and of Harlen Bretz (one of my own heroes) who was belatedly awarded the Penrose Medal at the age of 97. Yet there is also a great deal about a vast variety of geological topics, from teaching methods to the effects of current history on the geological world (though to a European the great depression and prohibition seem to figure larger than the two World Wars).

Included, too, are day-by-day accounts of Pettijohn's long early canoe journeys to the north, with painful portages and campsites on Indian graveyards, plus the ever present mosquitoes and blackflies. We experience with him the excitement of going and seeing geology, for instance the discovery of a Precambrian tillite and the three-billion-year-old conglomerate that was the subject of his PhD thesis (on a lake island in northern Ontario). And we are reminded of a world in the early 1930s when \$200 was enough for the expenses of a three-month field season on the Canadian Shield together with those of a graduate field assistant. Pettijohn was always mapping and studying sections, but he was at the same time asking fundamental questions such as how one produces pure quartz sands before the development of land vegetation.

Though the author's main interest is in the sedimentary rocks, he does not say much about fossils. One wonders if this is the result of an early education, in which the literal truth of the Bible was taken for granted, and a first teaching job in a small college where geology had started out as a "Department of Harmony of Science and Revelation". But among his many distinguished later students it is interesting to note John Scopes, the Tennessee school teacher who was to be tried in 1925 for teaching the theory of evolution.

It was observed of Winston Churchill that he both made history and wrote it. Much the same might be said of Francis Pettijohn. This is a splendid, readable history of American geology in the twentieth century, by one who played a great part in its making. □

Derek Ager is Professor in and Head of the Department of Geology at University College, Swansea.

Review supplements in *Nature*

Dates of the review supplements to be published in 1985 are: Textbooks 7 March; Spring Books 25 April; New Journals 26 September; and Autumn Books 14 November.

The journals review issue will cover journals which first appeared between June 1983 and May 1984. Publishers will be sent further details in May of next year.

The Pleistocene ways of death

Leonard Krishtalka

Quaternary Extinctions: A Prehistoric Revolution.

Edited by Paul S. Martin and Richard G. Klein.

University of Arizona Press: 1984.

Pp.892. \$65.

THE plot seems simple enough. At the end of the Pleistocene, between 12,000 and 10,000 years ago (yr BP), the mammalian faunas of two continents were decimated. North America lost 33 genera (or 75%) of its megafauna (body size of 44 kg or greater) and four genera of small mammals. The extinctions and extirpations struck two entire orders, the Proboscidea (mammoth, mastodons, gomphotheres) and Perissodactyla (horses, tapirs), the families of camels, ground sloths, glyptodonts and peccaries, and genera of cheetah, sabertooth cat, bears, giant rodents, deer, musk oxen and moose. In South America, the 46 genera that became extinct were all large mammals: edentates, rodents, carnivores, endemic ungulates (litopterns, notoungulates), horses, mastodonts, peccaries, camels and deer.

Australia's fauna fared no better. During the late Pleistocene, between approximately 40,000 and 15,000 yr BP, some 55 species vanished. The megafaunal tombstone reads two echidnas, two marsupial carnivores, three wombats, seven diprotodonts (large marsupial herbivores), 33 macropodids (kangaroos, wallabies and their relatives), a varanid lizard, a horned tortoise, three birds and a giant snake.

In contrast, the faunas of Europe and Africa emerged from the Pleistocene comparatively unscathed. Of the 13 genera that disappeared from Europe, three (woolly rhinoceros, woolly mammoth, giant deer) were true extinctions, whereas the remainder (rhinoceros, horse, dhole, hippopotamus, musk ox, hyaena, antelope, elephant) survived elsewhere. In Africa too there were numerous local extirpations, but only a handful of species (long-horned buffalo, giant hartebeest, giant Cape horse, two springbok and a warthog-like pig) vanished from the continent at the Pleistocene-Holocene boundary. Africa had already suffered the extinction of 21 genera (one primate, one carnivore, 19 artiodactyls) during the early Pleistocene (1.8-0.7 Myr BP), and nine genera (one elephant, one horse, seven artiodactyls) during the middle Pleistocene (0.7-0.13 Myr BP).

Asia, like Europe and Africa, appears to have suffered only modest losses at the end of the Pleistocene, but an accurate checklist of the genera and species involved has yet to be compiled.

Now the plot thickens: is it a "what-

dunit" or a "whodunit"? One school of palaeobiologists blames rapid climatic and environmental change, namely, the final glacial retreat at the end of the Pleistocene and its consequences: a decrease in equability and in the diversity, quality and quantity of plant resources; a concomitant increase in "continentality" of climate, in seasonal extremes, and in the homogeneity and zonation of plant communities. These biotic changes played havoc with late Pleistocene, co-evolved ecosystems, such as grazing and browsing associations and gestation and growth periods among mammalian herbivores, quickly leading to the extinction of the megafauna, as well as dwarfing among the extinct and surviving lineages. This "glacial retreat" model is specific to the Americas, Europe and Asia,

IMAGE UNAVAILABLE FOR COPYRIGHT REASONS

Peter Murray

Extinct — *Elasmotherium*, from the steppes of Eurasia.

but climatic-environmental change is also invoked to explain the earlier extinctions in Africa and Australia.

A second school implicates a different but familiar culprit, *Homo sapiens*. It claims that a rapidly spreading "front" of palaeolithic hunters decimated the now extinct megafauna of the Americas, beginning in Canada about 12,000 yr BP and ending in Patagonia some 2,000 years later. The same "overkill" occurred in Australia following man's first invasion about 40,000 yr BP; in Africa, the shadow of Acheulean culture hangs over the early Pleistocene extinctions.

For more than 100 years natural historians have camped on either side of this Pleistocene battlefield: climatic-environmental change versus human overkill. Into the fray now marches *Quaternary Extinctions: A Prehistoric Revolution*, a volume armed with 38 chapters and 47 contributors. Extinction is no laughing matter, but, after reading this work, I'm reminded of the quip of the American humorist, Robert Benchley: "The world is divided into two kinds of people — those who divide the world into two kinds of people and those who don't". The two camps, still divided, are almost as entrenched as ever. Nevertheless, the editors, Paul S. Martin and Richard G. Klein, have produced a superb single-volume treatment of Pleistocene-Holocene extinctions and of the debate which surrounds them.

The 38 chapters are organized into seven

sections. The first sets the stage with papers on nineteenth-century explanations of Pleistocene extinctions, a bestiary of Pleistocene mammals and an account of New World mammoth distribution. Section 2 examines five significant late-Pleistocene sites, four in North America and one in Venezuela. Sections 3 and 4 present "the theoretical marketplace" through seven papers on the climatic-environmental version of extinction and six on the overkill model. These contributions deal mostly with the North American record, but Paul Martin's chapter, the longest and perhaps most eloquent in the book, presents the case for overkill on a global scale. Also, he alone deals in detail with the extinct Pleistocene faunas of South America and Europe. Section 5 examines the record of extinctions in northern Eurasia, China (mammoth only), the Levant, Africa and Madagascar, while the sixth covers environmental change, the archaeological record and the severe extinctions in Australia, New Zealand and the Hawaiian Islands.

Taken together, these first 35 chapters comprise an indispensable and almost encyclopaedic reference text on Pleistocene and Holocene faunas, chronologies, environments and extinctions. The data, analyses and bibliographies are massive in amount, virtually worldwide in scope and up to date. But the volume doesn't end there.

In the seventh and final section, an excellent and incisive three-part overview, L.G.

IMAGE UNAVAILABLE FOR COPYRIGHT REASONS

Extinct — *Macrauchenia*, a litoptern from South America.

Marshall, D.K. Grayson and J.L. Diamond step back from the myriad facts, figures and interpretations to view matters in a broader, more comprehensive light. The particular illumination provided by Marshall is that of conceptual clarity. Ten issues — terminology, scientific method, chronology/correlation/causation, historical perspective, extinction, turnover, megafauna, dwarfing, overkill and climate — and their treatment by the various authors are dissected, compared and evaluated. For Grayson, the history and structure of the debate on Pleistocene extinctions is the barometer by which to judge the competing environmental and overkill schools. Diamond plumbs historic extinctions for the ecological and biogeographical principles and patterns that

could make sense of Pleistocene-Holocene events.

All three authors would postpone making the final verdict on the climate-overkill debate. I agree. Its resolution, despite the mass of data collected together in *Quaternary Extinctions*, awaits a more precise and dense palaeontological and archaeological record. Both models are rooted in exact chronological correlations. Crucial to the overkill model is whether the late Pleistocene extinctions indeed followed the first appearance and spread of a new predator, *H. sapiens*, in the Americas and Australia. As Grayson points out, this correlation is virtually the only evidence the overkill camp can marshal. Their selective acceptance of only the "good" dates — those that fit the model (for example dates for human beings in North America no older than 12,000 yr BP, and those for mammoths no younger than 10,000 yr BP) — may play fast and loose with the evidence that doesn't fit.

Other problems with the overkill theory are the *ad hoc* escapes from attempts to falsify the hypothesis. If overkill occurred in the Americas and Australia, why is there a dearth of kill sites in the archaeological record? Overkill partisans answer with the "blitzkrieg" addendum: overkill was so swift and devastating, that few kill sites are to be expected. Not only is this untestable, but contrary reasoning is applied to New Zealand, where abundant associations of prehistoric man and 22 extinct species of moas are cited as evidence for overkill. Moreover, one simulation of the blitzkrieg model for North America (Whittington and Dyke, Chapter 20) allows for overkill to have occurred in as few as 250 or as many as 22,000 years. This, as Grayson notes, "removes the critical chronological test for the general overkill hypothesis" (p. 812). No one doubts the hand of man in the extinction of faunas of New Zealand, Madagascar, Hawaii and other islands within the past 1,000 years, but, as Diamond and others show, overkill was merely one of a litany of lethal weapons, including habitat destruction and introduction of herbivores, predators, competitors and disease. In many instances, the effects of climatic change and man added up to a one-two punch.

Does the climatic model fare better? Somewhat. It too owes its first allegiance to chronological correlation, namely, the synchronicity of megafaunal disappearances and terminal Pleistocene post-glacial climatic change. But, as with the overkill model, a correlation is not a cause; it is merely a good reason to look deeper for the knots that tie climate to extinction, specifically, cause-effect processes and evidence that link the degree and kind of climatic, environmental and ecological change to observed extirpations and extinctions. Here, preliminary research has been encouraging. According to palaeobotanical evidence cited by a few of the authors, the breakdown of climatic

equability at the onset of the last interglacial may have been much more severe than in previous interglacials. If so, this explains why the megafauna would be extinguished during the last glacial retreat after having survived at least three previous ones. Five of the chapters begin to implicate post-glacial vegetational zonation and simplification in the North American extinctions, and one paper indicts late

IMAGE UNAVAILABLE FOR COPYRIGHT REASONS

Extinct — *Nothrotheriops*, the Shasta ground sloth from North America.

Pleistocene aridity cycles for the Australian disappearances. Unlike the overkill model, much of the climatic-environmental case is testable in the fossil record.

But the climatic model too has problems. Diamond wonders why the megafaunal extinctions were not accompanied by equivalent plant extinctions, how poikilotherms (reptiles, amphibians, fish and beetles, for example) escaped devastation when homoeotherms (mammals) did not, and why the final glacial retreat ravaged the North American megafauna but barely scarred the Eurasian one. Beyond the need for more data, the answers to these questions and, ultimately, to the overkill-climate debate, await a less polarized approach to the problem. Different causes of extinction may apply to particular faunas, landmasses and their subsets. Similarly, a one-two effect of climatic-environmental change and cultural decimation — acting in different proportions — may account for the selective and differential extinctions between landmasses and faunas. Most of the workers in the climatic camp have begun to think along these lines.

Two flaws in *Quaternary Extinctions* deserve mention. Except for Martin's chapter, South America and Europe receive little treatment and Asia virtually none. This may be more a comment on the state of the science than an act of omission. More minor a failing, but irritating nonetheless, is the absence of a list of institutional affiliations of the authors.

This work is a welcome sequel to Martin and Wright's *Pleistocene Extinctions: The Search for a Cause* (Yale University Press, 1967), which is out of print and out of date. As such, *Quaternary Extinctions* represents a leap in our knowledge of Pleistocene and Holocene palaeobiology. Many volumes on our bookshelves are destined to gather dust rather than attention. But not this one. □

Leonard Krishtalka is Associate Curator in the Section of Vertebrate Fossils, Carnegie Museum of Natural History, Pittsburgh, Pennsylvania.

Reductions in ozone at high concentrations of stratospheric halogens

Michael J. Prather, Michael B. McElroy & Steven C. Wofsy

Center for Earth and Planetary Physics, Harvard University, Cambridge, Massachusetts 02138, USA

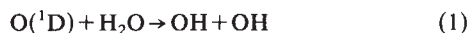
An increase in the concentration of inorganic chlorine to levels comparable to that of oxidized reactive nitrogen could cause a significant change in the chemistry of the lower stratosphere leading to a reduction potentially larger than 15% in the column density of ozone. This could occur, for example by the middle of the next century, if emissions of man-made chlorocarbons were to grow at a rate of 3% per year. Ozone could be further depressed by release of industrial bromocarbon.

OUR understanding of the stratosphere has grown steadily since Chapman's¹ original discussion of photochemical processes in a pure oxygen system. Milestones include the introduction of reactions involving hydrogen compounds², recognition of the importance of oxides of nitrogen^{3,4} and, more recently, realization that a complete model must allow for reactions involving chlorine⁵ and bromine^{6,7}. Current models for stratospheric chemistry include as many as 200 reactions coupling various forms of oxygen, hydrogen, nitrogen, chlorine and bromine.

The emphasis here is on chemical processes in the lower stratosphere, between 20 and 30 km. Much of the work in recent years has been concerned with assessing the effects of various forms of anthropogenic activity on the stability of the ozone column. Emissions of chlorocarbons are thought to influence ozone (O_3) primarily in the region above 30 km. We shall argue that important reductions in $[O_3]$ could take place also below 30 km if the concentration of chlorine were to rise above that of NO_x (defined as the composite of all oxidized forms of nitrogen, excluding N_2O). Changes below 30 km might be expected to lead to an abrupt drop in the column density of O_3 with increasing levels of chlorine, in association with a marked change in the chemical characteristics of the lower stratosphere.

Chemistry of the stratosphere

The abundance of OH in the lower stratosphere is at present controlled mainly by:



followed by



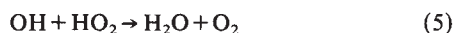
and



Nitric acid is the most abundant form of NO_x , followed by NO_2 , N_2O_5 and HNO_4 . At higher levels of chlorine, $CINO_3$, formed by



is more important. Concentrations of NO_2 , HNO_3 and HNO_4 decrease rapidly as $CINO_3$ becomes the dominant form of NO_x . Removal of OH is then effected by the less efficient loss mechanisms:



and



with additional production of odd-hydrogen due to oxidation of CH_3 , formed by



The concentration of OH increases rapidly at higher levels of chlorine, as illustrated in Fig. 1, owing in part to titration of NO_x to $CINO_3$, in part to enhanced production of $O(^1D)$ caused by deeper penetration of ultraviolet radiation resulting from reduction in the overlying column of O_3 . Similar shifts in the chemistry of OH and NO_x were noted earlier by Prather *et al.*⁸, who emphasized the dramatic alterations in radical chemistry which occur when NO and NO_2 are converted to $CINO_3$. They raised the possibility of multiple steady states for OH in the high latitude winter stratosphere, under conditions where inorganic chlorine ($[Cl_x] = [Cl] + [ClO] + [CINO_3] + [HOCl] + [HCl]$) exceeds $[NO_x]$.

It is convenient to identify a family of species, including O_3 and other compounds, which convert rapidly to O_3 in the lower

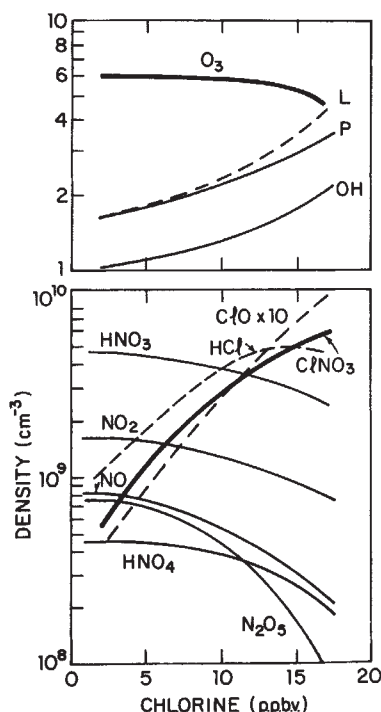


Fig. 1 Changes in the chemistry of the stratosphere at 24 km as a function of total chlorine concentration. Model calculations allowed for the diurnal variations in sunlight and chemical species, as described in the text. Inorganic chlorine (HCl , Cl , ClO , $CINO_3$, $HOCl$) accounts for approximately 70% of the total chlorine at 24 km, increasing to 100% by 40 km. The top panel shows a 20% decrease in $[O_3]$ (10^{12} cm^{-3}) as chlorine increases from 2 to 16 parts per 10^9 . Also shown are increases in $[OH]$ (10^8 cm^{-3}), in production of odd-oxygen (P , $2 \times 10^5 \text{ cm}^{-3} \text{ s}^{-1}$) and in the loss frequency for odd-oxygen (L , yr^{-1}). The lower panel illustrates a rapid rise in $[ClO]$ and an accompanying change in partitioning of NO_x as $CINO_3$ becomes the dominant form of NO_x .

Table 1 Reaction sequences involving odd-oxygen

Production of odd-oxygen	
(0)	$O_2 + h\nu \rightarrow O + O$
Loss of odd-oxygen	
(I)	$O + O_3 \rightarrow O_2 + O_2$
(II)	$HO_2 + O_3 \rightarrow OH + O_2 + O_2$ $OH + O_3 \rightarrow HO_2 + O_2$
	$2O_3 \rightarrow 3O_2$
(III)	$NO_2 + O \rightarrow NO + O_2$ $NO + O_3 \rightarrow NO_2 + O_2$
	$O + O_3 \rightarrow 2O_2$
(IV)	$ClO + O \rightarrow Cl + O_2$ $Cl + O_3 \rightarrow ClO + O_2$
	$O + O_3 \rightarrow 2O_2$
(V)	$BrO + O \rightarrow Br + O_2$ $Br + O_3 \rightarrow BrO + O_2$
	$O + O_3 \rightarrow 2O_2$
(VI)	$ClO + BrO \rightarrow Cl + Br + O_2$ $Cl + O_3 \rightarrow ClO + O_2$ $Br + O_3 \rightarrow BrO + O_2$
	$2O_3 \rightarrow 3O_2$
(VII)	$ClO + HO_2 \rightarrow HOCl + O_2$ $HOCl + h\nu \rightarrow OH + Cl$ $OH + O_3 \rightarrow HO_2 + O_2$ $Cl + O_3 \rightarrow ClO + O_2$
	$2O_3 \rightarrow 3O_2$
(VIII)	$ClO + NO_2 + M \rightarrow ClONO_2 + M$ $ClONO_2 + h\nu \rightarrow Cl + NO_3$ $Cl + O_3 \rightarrow ClO + O_2$ $NO_3 + h\nu \rightarrow NO + O_2$ $NO + O_3 \rightarrow NO_2 + O_2$
	$2O_3 \rightarrow 3O_2$
No change in odd-oxygen	
(IX)	$HO_2 + NO \rightarrow OH + NO_2$ $OH + O_3 \rightarrow HO_2 + O_2$ $NO_2 + h\nu \rightarrow NO + O$
	$O_3 \rightarrow O + O_2$
(X)	$ClO + NO_2 + M \rightarrow ClONO_2 + M$ $ClONO_2 + h\nu \rightarrow Cl + NO_3$ $Cl + O_3 \rightarrow ClO + O_2$ $NO_3 + h\nu \rightarrow NO_2 + O$
	$O_3 \rightarrow O + O_2$

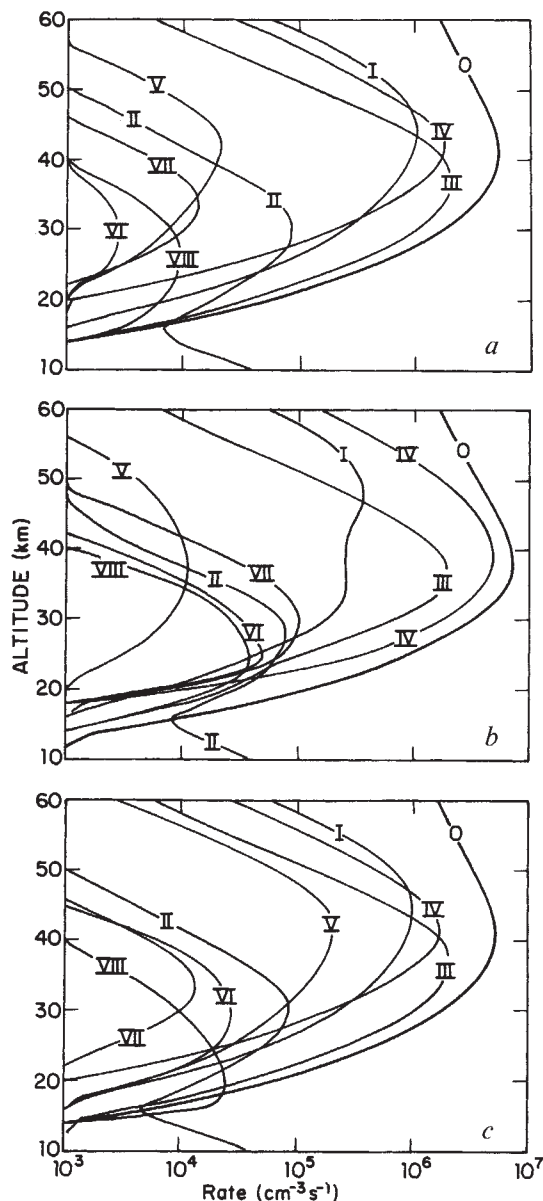
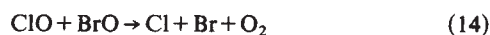
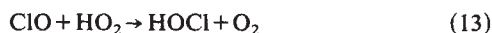
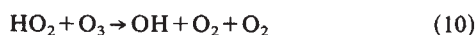


Fig. 2 Vertical profiles of diurnally averaged rates for production and loss of odd-oxygen. The rates (odd-oxygen equivalents $\text{cm}^{-3} \text{s}^{-1}$) are calculated with the model described in the text, including the diurnal variation in sunlight. The curves are labelled by their reaction sequence as given in Table 1: for example, curve III represents twice the rate for reaction (11), averaged over 24 hours. *a*, Our best estimate of the contemporary atmosphere (2.6 parts per 10^9 chlorine); *b*, a perturbed environment with 16.4 parts per 10^9 chlorine; *c*, the effect of bromine at a concentration of 200 parts per 10^{12} . The reactions $O + OH$ and $O + HO_2$, which dominate odd-oxygen loss above 48 km, are not shown. Curve VIII includes all photolysis of NO_3 which produces $NO + O_2$.

stratosphere. We term this family 'odd-oxygen': O , O_3 , HO_2 , NO_2 , ClO and BrO . Odd-oxygen is produced by photolysis of O_2 in the Herzberg continuum. It is removed by a variety of reaction sequences summarized in Table 1. The primary sinks are:



and

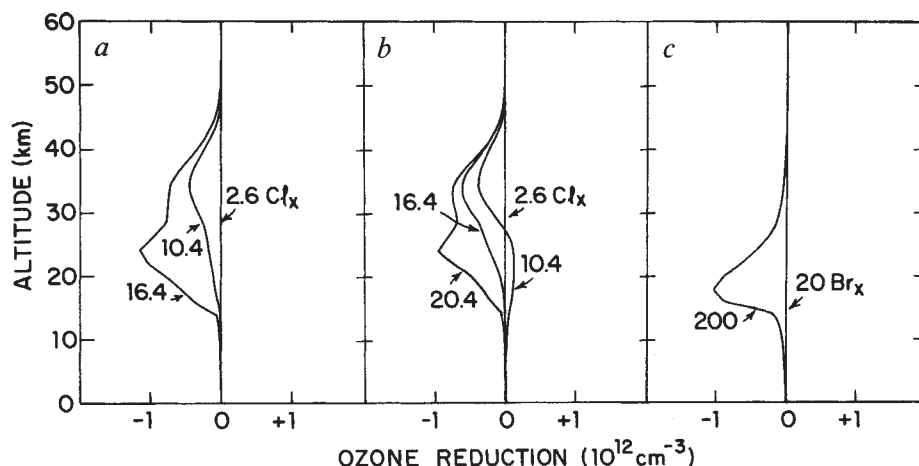


Each of these reactions accounts for removal of two odd-oxygen equivalents. Formation of $ClONO_2$ provides a temporary sink for odd-oxygen. This is balanced mostly by release of odd-oxygen in the photolysis of $ClONO_2$ and NO_3 (Table 1: sequence X). Reactions (9)–(11) account for most of the odd-oxygen removal below 30 km in the present atmosphere; reactions (12)–(14) dominate at high levels of chlorine.

Effects of chlorine

Vertical profiles for O_3 were calculated using a one-dimensional model⁹ with diffusion coefficients calibrated to match observational data for $CFCl_3$, CF_2Cl_2 , N_2O and CH_4 (ref. 10). Rates

Fig. 3 Vertical profiles of perturbations to the concentration of ozone for different levels of chlorine (Cl_x) and bromine (Br_x). The model allows for diurnal variations in sunlight. Panels *a* and *c* use our best estimate for the contemporary atmosphere (2.6 parts per 10^9 chlorine, 13.0 parts per 10^9 NO_x , 20 parts per 10^{12} bromine) as a reference profile for ozone; *b* considers a reference value with higher $[\text{NO}_x]$ (16.2 parts per 10^9). Panels *a* and *b* examine the effects of additional chlorine at levels of 10.4, 16.4 and 20.4 parts per 10^9 . Panel *c* shows the calculated change in ozone if bromine were to increase from 20 to 200 parts per 10^{12} .



for chemical reactions were taken from the National Aeronautics and Space Administration (NASA)¹¹, except that for (12) for which we used more recent measurements by Leu¹². Cross-sections for photolysis of O_2 are consistent with observations by Herman and Mentall¹³ and Anderson and Hall¹⁴ ($7.5 \times 10^{-24} \text{ cm}^2$ at 205 nm). Solar declination was set equal to 0° (equinox), and the temperature structure of the atmosphere follows the 30° N summer model of the US Standard Atmosphere¹⁵. Some calculations were carried out using rates for photolysis averaged over a 24-hour cycle, whereas others allowed for diurnal variations in sunlight. Both approaches have been used in previous investigations of the stratosphere⁹. Concentrations of CH_4 , N_2O , NO_x , Cl_x , Br_x and halocarbons were calculated self-consistently, for the present environment, using the one-dimensional formulation. The model for higher concentrations of chlorine was based on an assumption that growth in $[\text{Cl}_x]$ is caused primarily by enhanced emission of CF_2Cl_2 and CFCl_3 . A variety of exploratory calculations were also carried out, in which stratospheric profiles of CH_4 , NO_x and Br_x were scaled by constant factors, to explore the sensitivity of model results to possible changes in the concentrations of these species.

Rates for removal of odd-oxygen averaged over a 24-hour cycle, obtained with the more complete model allowing for the diurnal variation in sunlight, are illustrated in Fig. 2. Figure 2a shows results for the present stratosphere where the chlorine abundance is taken as 2.6 parts per 10^9 —significantly higher than the natural background, which is expected to be less than 1 part per 10^9 ; note that reaction (12) (Table 1: sequence IV) is the principal loss process for odd-oxygen between 40 and 45 km in the present atmosphere; reaction (11) (sequence III) is dominant below this. Results with the chlorine abundance equal to 16.4 parts per 10^9 are given in Fig. 2b: reaction (12) dominates over an extensive height range from 20 to 55 km; sequences II, III, VI, VII and VIII are comparably important, each contributing $\sim 10\%$ to the total loss of odd-oxygen below 25 km.

The enhanced role of reaction (12) in Fig. 2b reflects the much larger abundance of ClO throughout the stratosphere. The abundance of total chlorine in Fig. 2b exceeds that in Fig. 2a by a factor of 6; however, the abundance of ClO at 24 km is larger by a factor of 18. The non-linearity in $[\text{ClO}]$ results in part from the large increase in $[\text{OH}]$ which arises at high concentrations of chlorine. As $[\text{NO}]$ decreases, further enhancement of $[\text{ClO}]$ occurs owing to significant reduction in the rate at which ClO is cycled to Cl and HCl by



and reaction (8).

Changes in the vertical distribution of ozone as a function of increasing chlorine concentration are shown in Fig. 3: Fig. 3a shows model calculations for 13 parts per 10^9 NO_x at 35 km—our

best estimate of the present situation; Fig. 3b considers $[\text{NO}_x]$ equal to 16.2 parts per 10^9 , near the upper range of present observational data. The maximum relative depletion in $[\text{O}_3]$ occurs near 40 km, and $[\text{O}_3]$ is a monotonic and nearly linear function of added chlorine at this altitude. The behaviour in the lower stratosphere is quite different: at the higher $[\text{NO}_x]$ considered in Fig. 3b, addition of chlorine leads initially to a small increase in $[\text{O}_3]$ centred near 24 km; for chlorine concentrations above ~ 12 parts per 10^9 , the increase disappears and, at higher levels of chlorine, reduction in $[\text{O}_3]$ is observed at all altitudes. Conversion of NO_x to ClNO_2 requires less chlorine for the lower $[\text{NO}_x]$ case modelled in Fig. 3a. Large reductions in $[\text{O}_3]$ accordingly occur at smaller concentrations of chlorine, accounting for differences in the behaviour exhibited by the two cases in Fig. 3. These reductions arise despite significant enhancements ($\times 2$) in the production of odd-oxygen as illustrated in Fig. 1.

Figure 4a shows column densities of ozone for the cases summarized in Figs 2 and 3. Major reductions, $>10\%$, occur for both the high and low $[\text{NO}_x]$ cases when chlorine exceeds 18 parts per 10^9 . Further additions of chlorine, as little as 2 parts per 10^9 , can lead to a doubling of the resulting perturbation in ozone concentration.

The results in Figs 2, 3 and 4a account for the diurnal variations in sunlight and chemical species. Results in Fig. 4b–e explore the sensitivity of predicted ozone depletions to $[\text{NO}_x]$, $[\text{CH}_4]$ and other parameters of the model, using the computationally easier approach in which sunlight is averaged over a diurnal cycle. Use of an average flux to simulate the proper diurnal variation in sunlight leads to an overestimate for $[\text{NO}_2]$ and consequently $[\text{ClNO}_2]$. This accounts for the quantitative differences between the results in Fig. 4a and b: column $[\text{O}_3]$ is higher in Fig. 4a and the onset of non-linearity is apparent at smaller values of $[\text{Cl}_x]$.

Figure 4b illustrates the response of O_3 to chlorine, for $[\text{NO}_x]$ equal to 13.0, 16.2 and 19.5 parts per 10^9 at 35 km. The concentration of N_2O , the source of stratospheric NO_x , is increasing now by between 0.2 and 0.4% per year¹⁶. The higher levels of NO_x envisaged here, 16.2 and 19.5 parts per 10^9 , would correspond to increases in $[\text{N}_2\text{O}]$ of 25% and 50% respectively, relative to the present value of about 300 parts per 10^9 . The column density of O_3 changes little with addition of chlorine, so long as the concentration of chlorine is significantly less than that of NO_x .

Cicerone *et al.*¹⁷ have reported evidence for nonlinear behaviour in total $[\text{O}_3]$ as a function of chlorine concentration for values of $[\text{Cl}_x]$ below 10 parts per 10^9 . This behaviour was also observed in the present study. Nonlinearities may arise at low chlorine concentrations owing to compensatory effects in the photochemistry of the lower stratosphere, enhancing production but also removal of odd-oxygen. However, the net impact on $[\text{O}_3]$ is small, and results are quantitatively sensitive to details of the underlying model. The nonlinearity at high chlorine

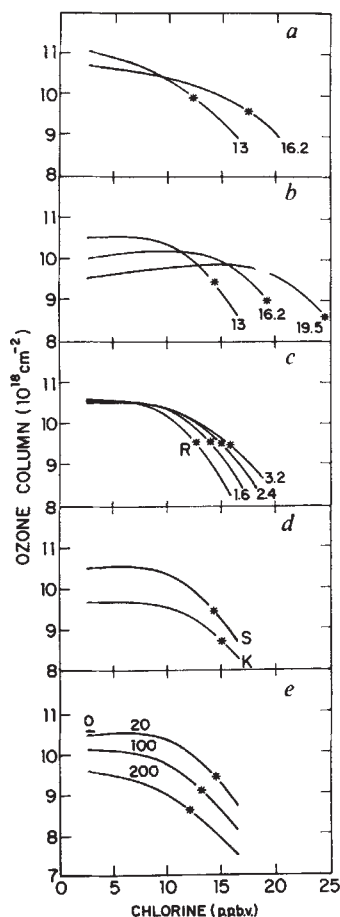


Fig. 4 The column abundance of ozone as a function of stratospheric chlorine. Results in panel *a* allow for diurnal variations in sunlight. The sunlight was averaged over 24 hours for *b*–*e*. Panel *a* shows results for our standard contemporary atmosphere (13 parts per 10^9 NO_x at 35 km) and for an atmosphere with 25% more NO_x (16.2 parts per 10^9). Panel *b* includes a case with 50% higher $[\text{NO}_x]$ (19.5 parts per 10^9). Asterisks denote the point at which the ozone column has been depleted by 10% relative to the contemporary atmosphere (2.6 parts per 10^9 chlorine). Panel *c* considers effects of an increase in $[\text{CH}_4]$ from 1.6 p.p.m. to 2.4 p.p.m. to 3.2 p.p.m. *c* also includes a case *R* in which the rate coefficient for reaction (6) was changed to $2.7 \times 10^{-12} \exp(-360/T)$; this reflects recent kinetic data (L. F. Kaiser, unpublished; ref. 36; see also ref. 17), indicating that this reaction may be 30% faster than previously suggested. Panel *d* compares ozone perturbations calculated for the standard model *S* with those for a model *K* with slow vertical diffusion (see text). Panel *e* presents results from calculations with bromine concentrations of 0, 20, 100 and 200 parts per 10^{12} .

concentration is large and robust: it can be interpreted essentially as a local photochemical effect and should be apparent in all models where $[\text{Cl}_x]$ exceeds $[\text{NO}_x]$.

The concentration of atmospheric CH_4 appears to have risen by between 1 and 2% per year over the past several years^{18,19}, and perhaps by as much as a factor of 2, since the sixteenth century²⁰. An increase in $[\text{CH}_4]$ favours production of HCl through reaction (8), with consequent reduction in the ratio of $[\text{ClO}]$ to total chlorine concentration; it may be expected also to result in an enhanced concentration of stratospheric H_2O . The implications for ozone are summarized in Fig. 4c. Abundances for H_2O were calculated as functions of altitude, allowing for the CH_4 source, with the mixing ratio of H_2O fixed at 3 ppm at the tropopause. The major effect in Fig. 4c reflects the influence of CH_4 as a sink for reactive chlorine. At high $[\text{CH}_4]$, the sharp decline in $[\text{O}_3]$ occurs at higher concentrations of chlorine. For example, a 15% reduction in $[\text{O}_3]$ is predicted to

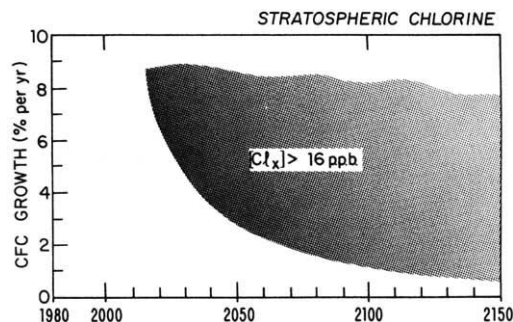


Fig. 5 The time at which the concentration of stratospheric chlorine is expected to exceed 16 parts per 10^9 (p.p.b.) as a function of growth in chlorofluorocarbon (CFC) emissions alone. The base calculation starts in 1980 with 2.6 parts per 10^9 chlorine, including 170 parts per 10^{12} of CFC-11, and 285 parts per 10^{12} of CFC-12. The initial emission rates were 264×10^9 g CFC-11 yr^{-1} and 422×10^9 g CFC-12 yr^{-1} and steady state lifetimes are 80 and 140 yr, respectively. The growth in emission rates was compounded annually, and the loss was calculated with a 2 yr lag to account for the delay in stratosphere-troposphere equilibrium. The abundance of chlorine in steady state would be 10.7 parts per 10^9 if emission rates were held constant at 1980 values. A doubling of emissions would provide an asymptotic abundance of 19.8 parts per 10^9 .

occur, in Fig. 4c, at a chlorine concentration of 16 parts per 10^9 , for present levels of CH_4 . A comparable reduction in $[\text{O}_3]$ would be postponed to a chlorine concentration of 18 parts per 10^9 , if the abundance of CH_4 was equal to twice its present value.

The results in Fig. 4a–c suggest that the average column density of O_3 should be $\sim 1.1 \times 10^{19}$ molecules cm^{-2} , or 410 Dobson units, for the present environment. This is higher than observed values, which are centred around 330 Dobson units (ref. 21). The discrepancy may reflect deficiencies in the one-dimensional model, in particular, its inability to account for the large spatial ($\times 2$) and temporal (50%) variability in ozone concentration. It could also reflect deficiencies in our understanding of the relevant chemistry, or the approach used to simulate vertical transport. Holton²² suggested that effective diffusion coefficients for chemically active gases are smaller than those for inert species. He offered a method whereby diffusion coefficients could be corrected to allow for effects of a finite chemical lifetime. Results in Fig. 4d incorporate this modification, setting the diffusion coefficient for O_3 equal to $3 \times 10^3 \text{ cm}^2 \text{ s}^{-1}$ throughout the lower stratosphere. The role of vertical transport is diminished accordingly. Photochemical processes are relatively more important for O_3 , and our estimate for the present-day column is reduced to approximately 370 Dobson units, in better agreement with observation. The response of O_3 to elevated chlorine concentration is similar, however, to the behaviour exhibited in the results given above.

The concentration of chlorine in the present stratosphere is set mainly by anthropogenic sources of CFCl_3 (20%), CF_2Cl_2 (20%), CCl_4 (20%) and CH_3CCl_3 (15%). Release rates for these gases are now about 260×10^9 , 420×10^9 , 120×10^9 and 500×10^9 g yr^{-1} , respectively^{23–25}. Associated lifetimes are 70, 130, 55 and 10 yr (refs 24–28). The concentration of stratospheric chlorine is expected to rise eventually to ~ 10 parts per 10^9 , if the release rate remains constant. This is less than the level required to induce the sharp reductions in $[\text{O}_3]$ below 30 km, and calculated perturbations to the ozone column are consequently small^{10,29}. Figure 5 indicates, as a function of the assumed growth rate, the year in which chlorine would be expected to rise to about 16 parts per 10^9 —our estimate for the threshold needed to induce a 15% reduction in $[\text{O}_3]$. If we focus, for example, on effects over the next 50 years, the results in Fig. 5 indicate that large reductions in $[\text{O}_3]$ are possible if the compounded growth of chlorofluorocarbon emissions exceeds 5% per year.

Effects of bromine

Emissions of bromine could also lead to a significant reduction in stratospheric ozone, according to reaction sequences V and VI (Table 1). A major fraction of bromine is present in the stratosphere in the reactive form BrO, and bromine is consequently more efficient than chlorine as a catalyst for removal of odd-oxygen. The concentration of organic bromine is now between 25 and 30 parts per 10^{12} (refs 30–32) with methylbromide (CH_3Br) accounting for 30 to 70% of the present atmospheric burden. The lifetime of CH_3Br is about 3 yr and known industrial sources account for approximately half of the current release³³. The gases ethylenedibromide ($\text{C}_2\text{H}_4\text{Br}_2$), 2.4 parts per 10^{12} as Br), and fluorocarbons FC-1301 (CF_3Br , 1 part per 10^{12}) and FC-1211 (CF_2ClBr , 1 part per 10^{12}) are mainly anthropogenic in origin and account for about 15% of the current bromine content of the atmosphere³².

The concentration of industrial bromocarbons is likely to rise in future, with the possible exception of ethylenedibromide which is derived mainly from use of leaded gasoline. Production of CH_3Br has increased four to fivefold since 1972 (ref. 34) and may be anticipated to rise further if it replaces ethylenedibromide as a fumigant. Use of FC-1301 and FC-1211 as fire extinguishers is growing, and concentrations of these gases in the atmosphere appear to have increased at a rate of $>10\%$ per year from 1978 to 1983 (R. A. Rasmussen and M. A. K. Khalil, personal communication, 1984). Residence times for FC-1301 and FC-1211 are long, about 110 yr and 25 yr respectively³⁵. Associated concentrations may therefore be expected to increase to levels of ~ 10 and 2.5 parts per 10^{12} in steady state, if release rates remain constant. The abundance of stratospheric bromine would rise eventually to about 100 parts per 10^{12} if emissions of brominated fire extinguishers were to grow to $20 \times 10^9 \text{ g yr}^{-1}$, approximately six times the present release. Allowing for possible growth in concentration of other brominated species, it seems appropriate to investigate the impact of a range of bromine concentrations, including values as much as an order of magnitude higher than those observed now.

Results for $[\text{O}_3]$ with bromine concentrations of 20 and 200 parts per 10^{12} are illustrated in Fig. 3c: maximum effects arise near 20 km; perturbations above 30 km are minimal. The absence of a significant bromine-related perturbation to $[\text{O}_3]$ in the upper stratosphere ensures that radiative feedbacks are relatively unimportant for bromine, in contrast to the situation for chlorine. Also, the concentration of bromine is much less than that of NO_x and formation of BrNO_3 provides an insignificant sink for NO_x . The response of the stratosphere is thus fairly linear with respect to addition of bromine.

The combined effects of bromine and chlorine are illustrated in Fig. 4e. Under present conditions for chlorine and NO_x , an increase in bromine from 20 to 100 parts per 10^{12} is calculated

to cause a 4% reduction in the column density of O_3 . Effects of chlorine and bromine are nearly additive. For a chlorine concentration of 16.4 parts per 10^9 , the reduction in $[\text{O}_3]$ is predicted to grow from 18% to 23%, as bromine is increased from 20 to 100 parts per 10^{12} . The results in Fig. 4e suggest that the column density of O_3 , in the present atmosphere, is $\sim 1\%$ less than it might otherwise be, in the absence of bromine.

Conclusions

In summary, models suggest modest, less than 3%, depletion in column O_3 if emission rates of chlorinated and brominated halocarbons remain constant. There are reasons, however, to believe that an increase in release rates could cause future problems for stratospheric O_3 . In particular, the chemistry of the lower stratosphere could undergo a significant change if the concentration of chlorine were to rise above that of NO_x . The concentration of OH could increase markedly and that of ozone could drop, with reductions potentially larger than 15% in the column density of O_3 . The O_3 column is expected to remain relatively constant before the onset of the NO_x to chlorine transition below 30 km. It is important not only to monitor the total column density of O_3 , but to search also for evidence of change in the chemical composition of the lower stratosphere. The large effects predicted here occur primarily below 30 km and involve major changes in the chemistry of radicals such as OH and ClO. Observational data for these species at lower altitudes are inadequate and yet are urgently required to test the validity of current models in a region containing the bulk of atmospheric ozone.

The present results are based on a one-dimensional model. They should be extended to consider variations of stratospheric chemistry with latitude and season. Extension of the model to more than one dimension may be particularly important in studying the impact of high concentrations of chlorine, as results should depend sensitively on details of the calculated latitudinal distributions of chlorine and NO_x .

Investigations of the radical chemistry of the stratosphere have in the past focused on the region above 30 km. It is clear that the emphasis should shift now to lower altitudes.

This research was supported by NASA grant NAGW-359 and by NSF grant ATM-81-17009.

Note added in proof: Preliminary independent studies (F. S. Rowland and H. Sato, personal communication; M. J. Molina, personal communication) suggest that HCl may react with ClONO_2 with subsequent release of Cl_2 . This could exacerbate the effects of the nonlinearities discussed here. An exploratory calculation indicates that reductions in column O_3 in excess of 15% could arise for chlorine concentrations as small as 10 parts per 10^9 if the associated rate for this reaction were as large as $10^{-16} \text{ cm}^3 \text{ s}^{-1}$.

Received 31 May; accepted 24 August 1984.

- Chapman, S. *Mem. R. met. Soc.* **3**, 103–125 (1930).
- Bates, D. R. & Nicolet, N. J. *geophys. Res.* **55**, 301–327 (1950).
- Crutzen, P. J. *Q. Jl R. met. Soc.* **96**, 320–325 (1970).
- Johnston, H. S. *Science* **173**, 517–522 (1971).
- Molina, M. J. & Rowland, F. S. *Nature* **249**, 810–812 (1974).
- Wofsy, S. C., McElroy, M. B. & Yung, Y. L. *Geophys. Res. Lett.* **2**, 215–218 (1975).
- Yung, Y. L., Pinto, J. P., Watson, R. T. & Sander, S. P. *J. atmos. Sci.* **37**, 339–353 (1980).
- Prather, M. J., McElroy, M. B., Wofsy, S. C. & Logan, J. A. *Geophys. Res. Lett.* **6**, 163–164 (1979).
- Logan, J. A., Prather, M. J., Wofsy, S. C. & McElroy, M. B. *Phil. Trans. R. Soc. A290*, 187–234 (1978).
- Causes and Effects of Changes in Stratospheric Ozone: Update 1983* (National Research Council, National Academy, Washington, D.C., 1984).
- Chemical Kinetics and Photochemical Data for Use in Stratospheric Modeling*, Evaluation No. 6 (JPL Publication 83-62, 1983).
- Leu, M. T. *J. phys. Chem.* **88**, 1394–1398 (1984).
- Herman, J. R. & Mentall, J. E. *J. geophys. Res.* **87**, 8697–8975 (1982).
- Anderson, G. P. & Hall, L. A. *J. geophys. Res.* **88**, 6801–6806 (1983).
- US Standard Atmosphere Supplements* (US Govt. Printing Office, Washington, D.C., 1966).
- Weiss, R. F. *J. geophys. Res.* **86**, 7185–7196 (1981).
- Cicerone, R. J., Walters, S. & Liu, S. C. *J. geophys. Res.* **88**, 3647–3661 (1983).

- Rasmussen, R. A. & Khalil, M. A. K. *Atmos. Envir.* **15**, 883–886 (1981).
- Ehhalt, D. H., Zandler, R. J. & Lamontagne, R. A. *J. geophys. Res.* **88**, 8442–8446 (1983).
- Craig, H. & Chou, C. C. *Geophys. Res. Lett.* **9**, 477–481 (1982).
- Hilsenrath, E., Heath, D. F. & Schlesinger, B. M. *J. geophys. Res.* **84**, 6969–6979 (1979).
- Holton, J. R. *J. geophys. Res.* (submitted).
- Production, Sales and Calculated Release of CFC-11 and CFC-12 Through 1982* (Chemical Manufacturers Association, Washington, D.C., 1983).
- Simmonds, P. G. *et al. J. geophys. Res.* **88**, 8427–8441 (1983).
- Prinn, R. G. *et al. J. geophys. Res.* **88**, 8415–8426 (1983).
- Ko, M. K. W. & Sze, N. D. *Nature* **297**, 317–319 (1982).
- Owens, A. J., Steed, J. M., Miller, C., Filkin, D. L. & Jesson, J. P. *Geophys. Res. Lett.* **9**, 700–703 (1982).
- Cunnold, D. M. *et al. J. geophys. Res.* **88**, 8379–8400 (1983).
- Present State of Knowledge of the Upper Atmosphere: An Assessment Report* (NASA, Washington, D.C., 1984).
- Singh, H. B., Salas, L. J. & Stiles, R. E. *J. geophys. Res.* **88**, 3684–3690 (1983).
- Berg, W. W., Sperry, P. D., Rahn, K. A. & Gladney, E. S. *J. geophys. Res.* **88**, 6719–6736 (1983).
- Rasmussen, R. A. & Khalil, M. A. K. *Geophys. Res. Lett.* **11**, 433–436 (1984).
- Bauer, E. A. *Catalog of Perturbing Influences on Stratospheric Ozone, 1955–1975* (Report FAA-EQ-78-20, Federal Aviation Administration, Washington, D.C., 1978).
- Minerals Yearbooks*, Vol. 1 (US Dept of Commerce, Washington, D.C., 1973–1981).
- Molina, L. T., Molina, M. J. & Rowland, F. S. *J. Phys. Chem.* **86**, 2672–2676 (1982).
- Molina, M. J., Molina, L. T. & Smith, C. A. *Int. J. chem. Kinet.* (in the press).

Microtubule assembly nucleated by isolated centrosomes

Tim Mitchison & Marc Kirschner

Department of Biochemistry and Biophysics, University of California at San Francisco, San Francisco, California 94143, USA

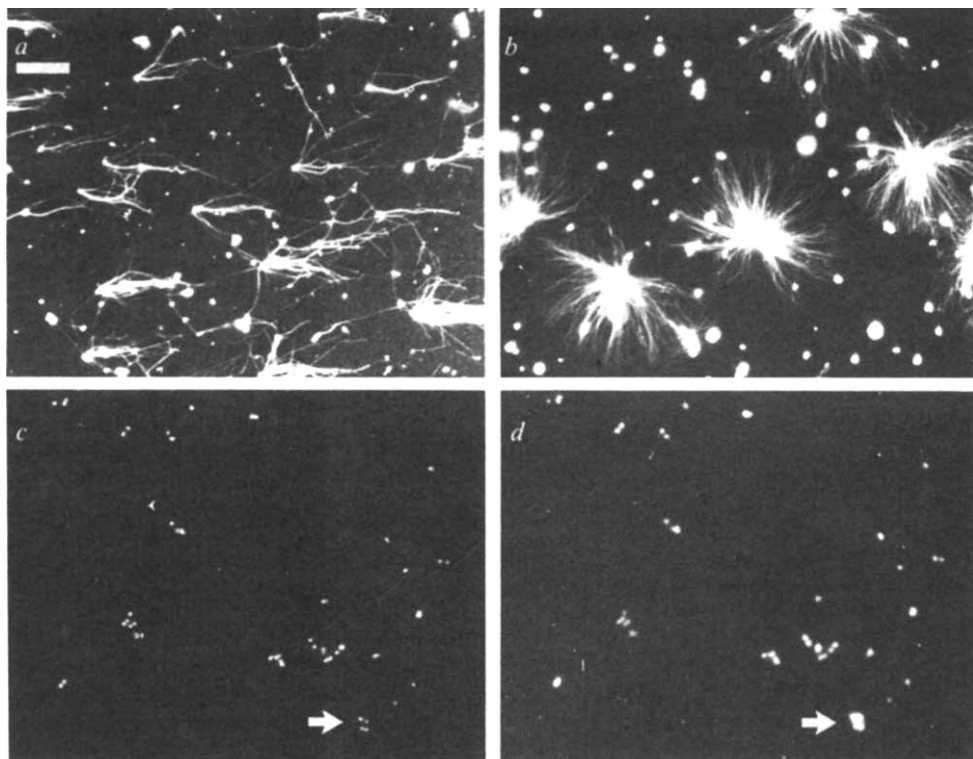
Microtubules are involved in the morphogenesis of most cells and are the structural basis of the mitotic spindle. We report here that purified centrosomes nucleate the assembly of microtubules with unusual dynamic properties. This may have important implications for the mechanism by which microtubule arrays are organized and stabilized in cells.

THE function of microtubules in cells is thought to depend on their specific spatial organization. In animal cells most microtubules are anchored at one end in a structure called the centrosome, which contains the centriole pair, and is surrounded by amorphous, osmiophilic material^{1,2}. In cells lacking centrioles, but containing a focus of microtubule growth, similar material is found, as are antigenic determinants common to those found in the centrosome, suggesting that there are common structural components in all interphase microtubule organizing sites³.

A fundamental question about the centrosome is how it organizes the assembly of microtubules *in vivo* and thereby, at least in part, determines the spatial arrangement of microtubules in the cell. This question can be broken into two parts, one concerning the nucleation, and the other the anchoring and stabilization of microtubules. Nucleation is essentially a kinetic process,

involving the coalescence of microtubule subunits into a seed, which can subsequently elongate. A structure which promotes nucleation will shorten or remove the lag phase in polymerization, and thus give microtubules attached to the structure a kinetic advantage. Microtubule nucleation by centrosomes is demonstrated *in vivo* when microtubules regrow preferentially from the centrosomes after cells are released from drug-induced microtubule depolymerization⁴⁻⁶. This aspect of centrosome function has also been demonstrated *in vitro*, using permeabilized cells^{7,8}, in crude lysates⁹⁻¹¹ and centrosome-nucleus complexes¹²⁻¹⁴. As well as giving local microtubules a kinetic advantage by nucleation, the centrosome *in vivo* must be able to anchor and preferentially stabilize them for long times. If centrosome microtubules had the same thermodynamic or steady-state stability as free microtubules, the latter eventually would come to

Fig. 1 Assay of centrosomes. **a**, Regrowth using centrosomes on coverslips. N115 centrosomes (see Fig. 2 legend) were diluted into 5 ml of PE buffer (10 mM PIPES, 1 mM EDTA, pH 7.2 with KOH) and sedimented onto 11-mm round glass coverslips at 25,000g (15 min, 4 °C) in a swinging bucket rotor. Corex tubes (15 ml) were modified to contain a plexiglass plug which could be lifted out easily with the coverslip on top. A drop of 5% Triton X-100 was added to each tube and the coverslip removed. 100 µl tubulin (35 µM in PB) was added to each coverslip, which was incubated at 37 °C for 12 min in a humidified chamber. Microtubules were fixed by aspirating most of the tubulin and adding 1% glutaraldehyde in PB (80 mM PIPES, 1 mM EDTA, 1 mM MgCl₂, 1 mM GTP, pH 6.8 with KOH) at 37 °C for 3 min. The coverslip was rinsed in PB' (PB without GTP) and postfixed in methanol at -20 °C. Immunofluorescence was with DMI α monoclonal anti-tubulin³⁴ and rhodamine-conjugated goat anti-mouse antibody (Cappel). Antibodies were routinely diluted in phosphate-buffered saline (PBS) + 1% bovine serum albumin (BSA) + 0.1% Triton X-100 + 0.05% NaN₃. Coverslips were washed in PBS + 0.1% Triton X-100 and mounted in 90% glycerol, 20 mM Tris, pH 7.8. Fluorescent microscopy was with a Zeiss Photomicroscope III, and photography used Kodak Technical Plan 2415 developed with HC110. **b**, Regrowth of centrosomes in solution. N115 centrosomes were mixed with tubulin at 0 °C in a final volume of 50 µl and a final concentration of 25 µM in PB. After 8 min at 37 °C, 200 µl of 1% glutaraldehyde in PB at 26 °C was added and then 3 min later, 1 ml of PB' at 0 °C. The regrown centrosomes were layered on top of a 5-ml cushion of 25% v/v glycerol in PB' in the modified corex tubes, and sedimented at 25,000g for 15 min onto a polysine-coated coverslip made by dipping a glass coverslip in 1 mg ml⁻¹ polylysine, then aspirating dry. The supernatant was aspirated, and the interface above the cushion washed with 1% Triton X-100. The coverslip was then removed, postfixed and stained as above. **c**, **d**, Structural assay of centrosomes using immunofluorescence. N115 centrosomes were sedimented on a plain glass coverslip as in **a**. Coverslips were fixed directly in methanol at -20 °C and stained with a mixture of monoclonal anti-tubulin antibody and 5051 human antipericentriolar material³⁵. Mixed rhodamine-conjugated goat anti-mouse and fluorescein-labelled goat anti-human (Cappel) secondary antibodies were used.



c, Tubulin; **d**, 5051. Scale bar corresponds to 6 µm in **a**, **b**; 3 µm in **c**, **d**.

predominate. Instead the reverse is found; free microtubules, if formed, are generally unstable with respect to centrosomal ones^{15,16}. This aspect of centrosome function has been discussed theoretically, and models have been proposed for centrosomal stabilization of microtubules by capping of a thermodynamically less stable end^{17,18} or by inducing a localized region in the cell which promotes microtubule assembly¹⁶. Experimental support for such models is at best indirect. To further study microtubule nucleation and stabilization, we have developed a method for the isolation of active centrosomes. By examining the assembly of pure tubulin onto isolated centrosomes, the kinetics of the nucleation process and the stability of nucleated microtubules was determined. The results of these studies have revealed important and unexpected features of microtubule polymerization, which bear on the role of the centrosomes in stabilizing and organizing microtubule arrays in cells.

Centrosomes and microtubule nucleation

To assay centrosomes during purification, we developed a functional assay in which microtubule assembly was dependent on centrosomes. To achieve this we needed a preparation of tubulin which polymerized well onto centrosomes but was deficient in spontaneous assembly. Tubulin purified by phosphocellulose chromatography in PIPES buffer without glycerol⁸ fulfilled these criteria and could be stored as aliquots at -70°C without subsequent change in assembly properties. This preparation of tubulin appeared to consist of only α - and β -tubulin on overloaded polyacrylamide gels, although the more sensitive technique of immunoblotting revealed residual τ -protein ($<0.1\%$ of the total protein present). Very little spontaneous assembly of this tubulin occurred at concentrations as high as $25\ \mu\text{M}$ ($2.5\ \text{g l}^{-1}$) for 30 min, yet it readily assembled onto centrosomes.

N115 neuroblastoma cells were initially chosen for centrosome isolation because they contain multiple centrioles and because the centrosomes seemed to be easily isolated in association with nuclei^{8,13,19-21} (Fig. 1). Centrosome preparations were exposed to tubulin and the extent of microtubule growth assayed by immunofluorescence. Spontaneous microtubule assembly was negligible. Centrosomes could be either regrown in solution or attached to a glass coverslip. The coverslip assay was more convenient in developing the purification scheme, as it was not limited either by centrosome concentration or by the buffer in which the centrosomes were placed. However, it was inferior for kinetic studies of centrosome-dependent microtubule growth, since the capacity of centrosomes was severely inhibited by adsorption to glass (Fig. 1a, b). For an accurate assessment of nucleation capacity, the solution assay was used. That the sites of microtubule regrowth were centrosomes was confirmed using antibody to tubulin, which recognizes the centriole cylinders (Fig. 1c), or with the human autoantibody against pericentriolar material (Fig. 1d)³. The staining pattern with these two antibodies gave similar distributions, showing the clustering of several centrioles into one N115 microtubule organizing centre. The pericentriolar material appears more irregular and diffuse, and sometimes connected adjacent centrioles (arrow in Fig. 1c, d). Immunofluorescence of naked centrosomes was used to determine their concentration during isolation, when conditions for preserving activity were already established. The antigenic determinants of both the centrioles and pericentriolar material were found to be much more robust than the nucleation capacity and are preserved by treatments which destroy nucleation capacity, such as 1 M NaCl or 2 M urea. Interestingly, treatment with 0.5 M KI or 2% sodium phosphotungstate at pH 7.0 appeared to solubilize the tubulin epitope (and destroy nucleation capacity) without perturbing the pericentriolar antigen staining. This suggests that the pericentriolar material may have a structural autonomy independent of the centriole cylinders.

We initially isolated complexes between centrosomes and nuclei from cells which were pretreated with nocodazole and cytochalasin B to depolymerize microtubules and weaken the actin meshwork¹³. However, there was a major difficulty in

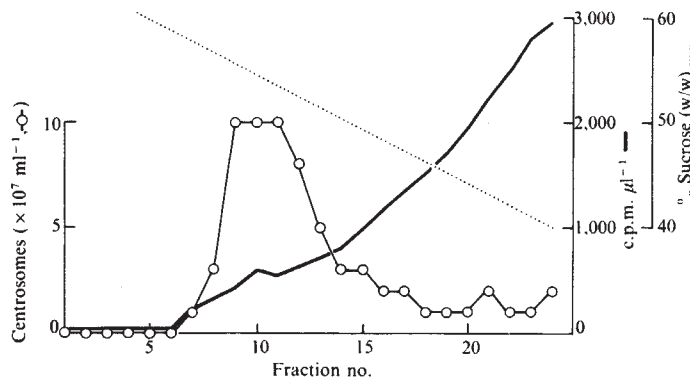


Fig. 2 Isolation of labelled N115 centrosomes by sucrose gradient sedimentation. Twenty confluent plates of N115 neuroblastoma cells were used. Cells were grown in DME H21 + 10% calf serum at 37°C in 7.4% CO_2 . For ^{35}S -methionine-labelled centrosomes, five plates of cells were labelled with 5 mCi of ^{35}S -methionine in medium containing 5% serum and 5% of the normal methionine complement for 18 h. These plates were combined with 15 unlabelled plates and processed as usual. Cells were pretreated with $10\ \mu\text{g ml}^{-1}$ nocodazole and $5\ \mu\text{g ml}^{-1}$ cytochalasin B in medium for 90–120 min at 37°C . Cells were then collected in medium and transferred to 50-ml polypropylene tubes, 10^8 cells per tube. All subsequent steps were at $0-4^{\circ}\text{C}$. Cells were washed using a table centrifuge at full speed ($1,500\text{g}$). Washes were with PBS, then PBS/10 + 8% sucrose, then 8% sucrose. The final pellet was resuspended in 12.5 ml per tube of 1 mM Tris-HCl, 0.1% βME (made using 2 M Tris-HCl stock, pH 8.0). To this was added 12.5 ml per tube of 1 mM Tris, 0.1% βME , 1% Nonidet P-40 (NP40). The lysate was spun at $1,500\text{g}$ for 3 min. The supernatant was filtered through $37\text{-}\mu\text{m}$ nylon mesh and 1/50 volume of $50\times\text{PE}$ (Fig. 1a) was added to the filtrate. This was transferred to 30-ml corex tubes and underlain with PE + 20% w/w Ficoll + 0.1% βME + 0.1% NP40, then spun at $25,000\text{g}$ for 15 min in a swinging bucket rotor at 2°C . Most of the supernatant was aspirated and the Ficoll interface collected with a Pasteur pipette (2.5 ml per tube). Ficoll interfaces were pooled and layered onto a sucrose gradient. The gradient was poured to half fill an SW28 tube (Beckman) and consisted of 20% (w/w)–62.5% (w/w) sucrose made up in PE + 0.1% βME + 0.1% Triton X-100. The gradient was spun at $27,000\text{ r.p.m.}$ for 1 h at 2°C in an SW28 rotor ($100,000\text{g}$); 30 drop fractions were collected from the bottom. To preserve nucleation capacity, the ionic strength was kept low throughout the preparation, but the activity was stable for several days at 0°C or several months at -70°C once in the final sucrose solution. Sucrose concentration was determined by refractometry, and centrosome concentration determined as for Fig. 1c, d. The three peak fractions were pooled, aliquoted, frozen on liquid nitrogen and stored at -70°C . To prepare centrosomes from CHO cells, the following modifications were made: washes were done on the tissue culture plate, using an aspirator to remove wash solutions; an extra wash of 1 mM Tris-HCl, 0.1% βME was used, then 10 ml per plate of the same buffer + 0.5% NP40 was added. Plates were rotated gently at 4°C for 10 min using a gyratory shaker (New Brunswick Scientific). The lysate containing centrosomes was collected and $50\times\text{PE}$ was added before the low-speed spin to remove chromatin. The filtration step was omitted and the preparation proceeded as above.

subsequently detaching the centrosomes from the nucleus, to which they are tightly associated during interphase^{9,14}. This proved very difficult, although treatment with 2 M urea in 100% D_2O buffers was partially successful in detaching the centrosomes and preserving nucleation activity. The attachment, however, appeared to be more labile in living cells, and we found that detergent lysis of cells at very low ionic strength liberated centrosomes into solution, while the same treatment of isolated nuclei did not. The low ionic strength lysis appeared to solubilize most of the cellular actin, giving a lysate containing little particulate material apart from some chromatin. For this treatment to work, it was essential for the metal cation concentration in the lysate to be $<0.5\ \text{mM}$. Centrosomes constituted the largest remaining particulate material after chromatin removal (Fig. 2).

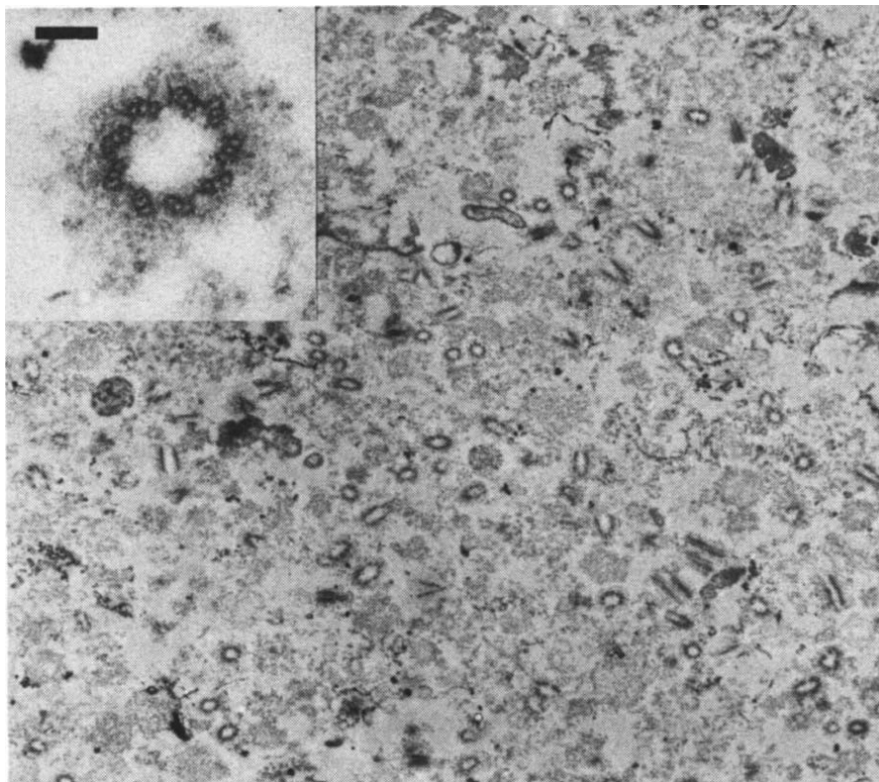


Fig. 3 Thin section of centrosomes. 10^8 N115 centrosomes, directly from the sucrose gradient, were diluted with 10 ml of 50 mM PIPES, 1 mM EDTA, pH 7.2 with KOH + 1% glutaraldehyde at 0°C and incubated for 30 min. The fixed centrosomes were then pelleted by spinning at 25,000g for 15 min. The pellet was washed twice in the same buffer without glutaraldehyde and resuspended in 2% OsO_4 in the same buffer. After 2 h at 0°C, the pellet was dehydrated through an ethanol series and propylene oxide, and embedded in Araldite. Sections (50 nm) were stained with uranyl acetate and lead citrate. Scale bar corresponds to 0.8 μm in the low-magnification field, and 0.1 μm in the inset.

Although N115 cells gave a good yield of centrosomes, they were heterogeneous in size, with one to ten centrioles per structure (Fig. 1c, d). As a source of more homogenous centrosomes, we used CHO cells which have a well characterized centriole cycle¹⁴. We modified the preparation slightly to take advantage of the adherent cell line and the more stable nucleus, but the effectiveness of a very similar protocol on two different lines suggests that it may have general use for centrosome purification (Fig. 2). The routine availability of a reproducible preparation of highly purified, stable centrosomes has opened the way for studies of centrosome function both *in vitro* and *in vivo*²².

Electron microscopy of a thin section of the N115 centrosome preparation reveals that the triplet structure of the centriole cylinder is well preserved and the pericentriolar material responsible for microtubule nucleation⁹ is visible (Fig. 3). We estimated that centriole cylinders constituted ~5% of the volume of the pellet. This is in agreement with quantitative immunoblotting and two-dimensional gel analysis, which indicated that 3–5% of the protein in the preparation was tubulin. When centrosomes were isolated from methionine-labelled cells (Fig. 2), the peak contained ~0.01% of the original counts; the overall yield obtained by counting centrioles was 30%, indicating a 3,000-fold purification. We estimate that the tubulin in a pair of centrioles should represent ~0.002% of the total cell protein, agreeing with our estimates that centrioles represent ~5% of the centrosome preparation. We cannot assess the functional importance of the remaining material, some of which must be the important pericentriolar components involved in microtubule nucleation.

Microtubule growth of centrosomes

Nucleation capacity of centrosomes was characterized by measuring the number and length of microtubules nucleated as a function of tubulin concentration. CHO centrosomes were regrown and fixed in solution, and then sedimented onto EM grids and visualized by rotary shadowing with platinum (Fig. 4b, c). Mean microtubule length increased at a constant rate at a given tubulin concentration, identical to the rate of increase for the plus ends of axonemes²³. This agrees with previous studies showing that *in vitro*, centrosomes nucleate microtubules with the plus ends distal¹⁰. We found considerable variation in microtubule number per centrosome, reflecting heterogeneity in

the population which may result from damage during isolation and the presence of different cell cycle stages. Mitotic centrosomes should be rare or absent in the preparation, since mitotic CHO cells would have been lost from the population during the plate washes. The number of microtubules nucleated per centrosome increased with tubulin concentration, starting at 4 μM and reaching a plateau above 20 μM (Fig. 4). This plateau may indicate saturation of nucleation sites, and the general shape of the curve is in agreement with earlier studies^{8,11,12}.

The average number of microtubules per centrosome increased gradually with tubulin concentration which is surprising as phase transitions generally show a step concentration dependence. This finding could be due to heterogeneity in the kinetics of nucleation, or to inherent differences in stability of microtubules grown at different concentrations. To examine this question, we measured the stability of free microtubules as a function of tubulin concentration.

To initiate the bulk polymerization of purified tubulin, we seeded its assembly with a small amount of microtubule fragments generated by spontaneous polymerization of the same tubulin in an assembly-promoting buffer containing 30% glycerol and 10 mM MgCl_2 ²⁴. Below 14 μM total tubulin, the microtubules depolymerize; above this concentration there is a linear increase of polymer with total tubulin, indicating a sharp phase transition (Fig. 5). The slope of this line is 0.68, probably indicating that two-thirds of the tubulin (stored as frozen aliquots) is competent for assembly. The concentration below which microtubules polymerize has previously been called the 'critical concentration'^{25,26}, by analogy with other phase transitions. We prefer the term 'steady state concentration' to indicate that this is the monomer concentration that exists when polymer is assembled to steady state¹⁸. When corrected for inactive tubulin, the true steady-state concentration is ~10 μM . However, we use here 14 μM , because this is the effective value for our preparation. This value may be compared with typical values found for tubulin and microtubule-associated proteins of 1–3 μM ^{27,28}. A 10-fold higher steady-state concentration for pure tubulin in aqueous PIPES buffer is in reasonable agreement with earlier estimates^{29,30}; pure tubulin may be incapable of polymerizing in less favourable aqueous buffers³¹.

There is a striking difference between the minimum concentration for microtubule assembly off centrosomes (3–4 μM) and

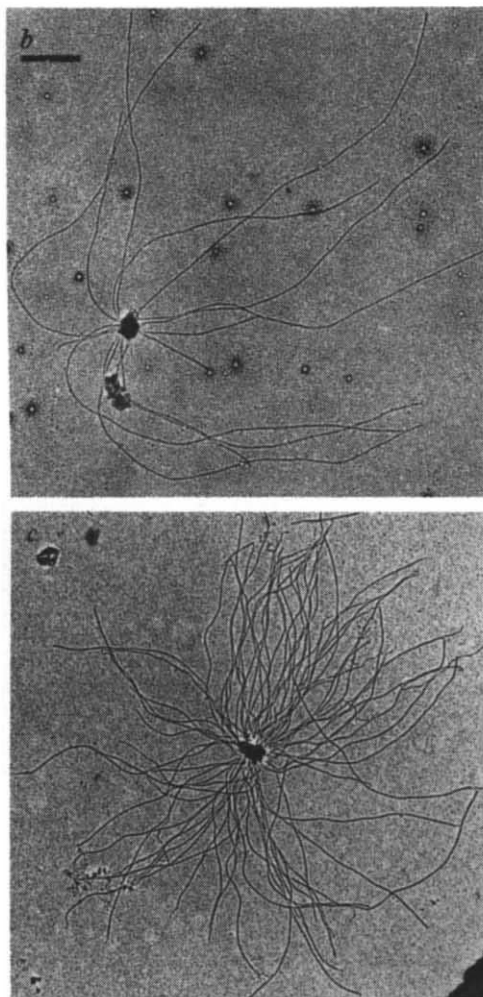
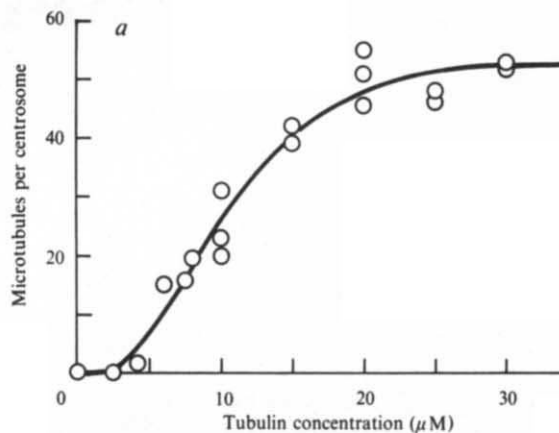


Fig. 4 Microtubule nucleation as a function of tubulin concentration. *a*, Number of microtubules nucleated per centrosome; *b*, *c*, typical images at 7.5 μM and 25 μM , respectively. Scale bar = 1.4 μm . Tubulin at appropriate concentration in PB was prewarmed for 2 min. 5×10^5 CHO centrosomes were added, and the mixture incubated in a final volume of 100 μl for 3–20 min at 37 $^\circ\text{C}$ (depending on tubulin concentration). The mixture was fixed by adding 300 μl of 1% glutaraldehyde in PB' at 26 $^\circ\text{C}$ and 3 min later, 600 μl of PB' at 0 $^\circ\text{C}$. In all operations involving pipetting or mixing regrown centrosomes, great care was taken to avoid shear by growing microtubules to equal length (10–15 μm) at different tubulin concentrations. 50 μl of the mixture were sedimented onto a 150-mesh Parlodion-coated grid using the EM90 rotor in the airfuge (Beckman) at 90,000g for 10 min. The grids were removed from the rotor, washed in 0.01% Triton X-100 and air-dried; they were then rotary-shadowed with platinum at an angle of 8 $^\circ$. Microtubules per centrosome were counted directly in the electron microscope at a magnification of $\times 5,000$. 100 centrosomes were counted per point.

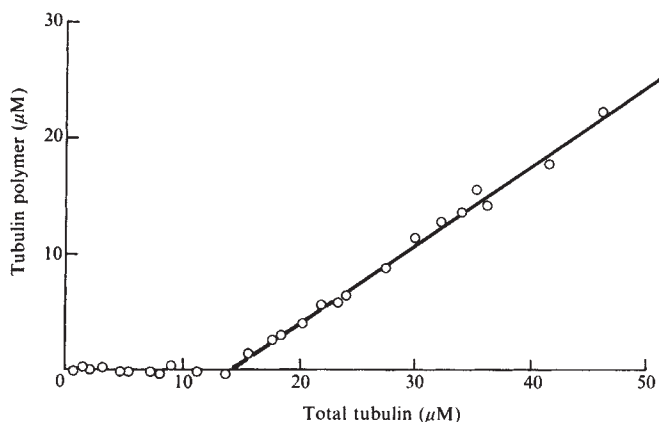


Fig. 5 Purification of tubulin and determination of steady-state concentration. 2-cycle beef brain microtubules²⁴ were resuspended and Dounce-homogenized in CB (50 mM PIPES, 1 mM EGTA, 0.2 mM MgCl_2 , pH 6.8 with KOH) and 1 mM GTP. After 20 min at 0 $^\circ\text{C}$ they were sedimented at 150,000g for 30 min at 2 $^\circ\text{C}$. 30 ml of the 10 mg ml^{-1} supernatant were applied to a 200-ml phosphocellulose column (Whatman P11) equilibrated in CB at 4 $^\circ\text{C}$. The column was eluted at 100 ml h^{-1} and the peak flow through fractions pooled; GTP was added to a concentration of 1 mM, and the tubulin was frozen in aliquots on liquid nitrogen and stored at -70 $^\circ\text{C}$. To induce spontaneous polymerization, tubulin was placed in a final buffer of 30% v/v glycerol, 80 mM PIPES, 1 mM EGTA, 10 mM MgCl_2 , 1 mM GTP pH 6.8 with KOH, at a tubulin concentration of 40 μM . After 40 min at 37 $^\circ\text{C}$, microtubules were sheared by six rapid passes through a 1.5-inch 22 gauge needle to produce uniform seeds. Tubulin was brought to a final buffer of PB (80 mM PIPES, 1 mM EGTA, 1 mM MgCl_2 , 1 mM GTP, pH 6.8 with KOH), and a GTP regenerating system (RS; final concentration 10 mM Na-acetyl phosphate + 1 U ml^{-1} of acetate kinase) was added³⁵. This solution containing 59 μM tubulin was prewarmed to 37 $^\circ\text{C}$, then 1/100 volume of seeds was added. After 30 min at 37 $^\circ\text{C}$, the solution was sheared by two passes as above. After 5 min at 37 $^\circ\text{C}$, the microtubule-containing solution was diluted by varying amounts, using warm PB + RS. Aliquots (200 μl) of diluted microtubules were allowed to equilibrate for 40 min at 37 $^\circ\text{C}$. Half of each aliquot was transferred to 0 $^\circ\text{C}$, and half was immediately sedimented at 90,000g for 4 min in the airfuge at 37 $^\circ\text{C}$. After 30 min at 0 $^\circ\text{C}$, the remaining half-aliquots were subjected to an identical spin at 4 $^\circ\text{C}$. Protein concentrations were determined by Coomassie blue binding assay using a bovine serum albumin standard and a molecular weight of 10^5 daltons for tubulin dimer. Polymer concentration was calculated as cold supernatant minus warm supernatant, and cold supernatant differed from total protein by <5%. The experiment was repeated on three tubulin preparations with similar results.

the steady-state concentration for assembly of pure tubulin (14 μM) (Figs 4, 5). Centrosomes seem able to nucleate microtubules well below the steady-state concentration, with microtubules first seen at 3–4 μM , despite the fact that free microtubules are unstable below 14 μM . Although this could possibly be due to differences in critical concentration between the two ends of the microtubule, we consider it unlikely as we have not found such differences in studying polymerization from axonemes²³. Furthermore, the amount of nucleation by centrosomes varies over a wide concentration range, whereas free microtubules show a sharp discontinuity in stability at the steady-state concentration and thus the expected shape of a phase transition (Figs 4, 5).

To study the effects of tubulin concentration on microtubule stability and avoid effects due to the kinetics of nucleation, intermediate tubulin concentrations were approached by diluting centrosomes presaturated with microtubules at high concentrations. Microtubule polymerization was initiated well above the steady-state concentration, resulting in centrosomes saturated with about 50 microtubules each. We found that if the tubulin was diluted to below the steady-state concentration, microtubules were lost from the centrosomes in a time- and concentration-dependent manner (Fig. 6). Surprisingly, the

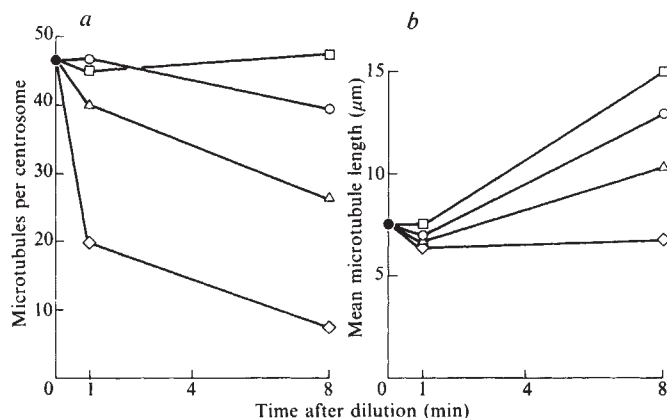


Fig. 6 Dilution of regrown centrosomes. *a*, Microtubules per centrosome; *b*, mean microtubule length. Microtubule nucleation was initiated at 25 μM tubulin as for Fig. 4. Regrown centrosomes were either fixed immediately, or diluted with warm PB, incubated further at 37 $^{\circ}\text{C}$ and fixed at the appropriate time (1–8 min). The final concentration of tubulin was 15 ($\square$), 10 ($\circ$), 7.5 ($\triangle$) or 5 μM ($\diamond$). Fixation and grid preparation were as described in Fig. 4 legend, except that the dilution was varied to keep the number of centrosomes per grid constant. To determine microtubule length, random centrosomes were photographed at a final magnification of $\times 3,000$. The negatives were digitized directly at $\sim \times 5$ magnification using a Numonics model 1224 digitizer interfaced to an HP85 computer. Absolute length was calibrated by photographing a calibration grid. 100 regrown centrosomes were counted and averaged for each point on *a* and 100 microtubule lengths averaged for each point on *b*. For numbers, s.d. was typically $\sim 30\%$ of the mean and for lengths at 8 min, s.d. was 2–3 μm . Thus, for sample sizes of 100, the number of microtubules at the lowest concentration and the length at the highest concentration differ from the initial values to a highly significant extent after 8 min.

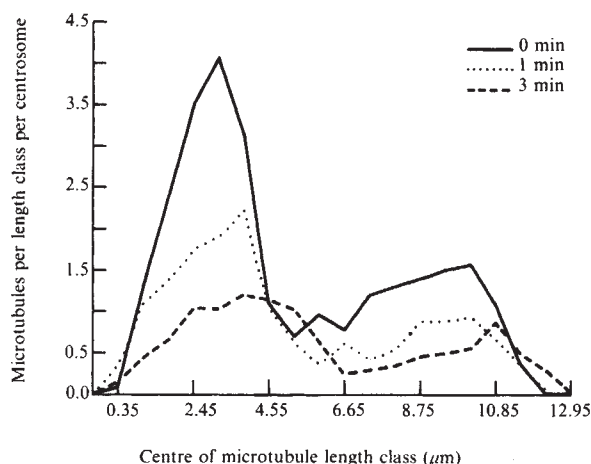


Fig. 7 Bimodal length experiment. Microtubule nucleation was initiated on centrosomes at 7.5 μM tubulin in PB. After 13 min at 37 $^{\circ}\text{C}$, sufficient prewarmed tubulin at 40 μM in PB was added to bring the solution to 25 μM . After 1.5 min at this concentration, an aliquot of the mixture was fixed and the remainder was diluted with warm PB to a final concentration of 5 μM . Aliquots were fixed 1 and 3 min after this final dilution. Fixed, regrown centrosomes were sedimented onto grids, photographed and measured as in Figs 5, 6. 100 centrosomes were counted to obtain the number per centrosome. For length distributions 700–1,000 microtubules were measured. The length distributions at each time were plotted as a histogram with 20 bars, then the ordinates of the histograms were scaled so that the y axis reflects the number of microtubules per centrosome in each size class. The lines connect the centre points of the top of each scaled histogram bar. The mean number of microtubules per centrosome was 27 (0 min), 16 (1 min) and 11 (3 min).

remaining microtubules continued to grow at a concentration-dependent rate. For example, when the centrosomes were diluted to 7.5 μM tubulin, the mean length increased by 40% ($P < 0.001$) but the mean number decreased by 40% ($P < 0.001$) in 8 min. This did not occur by the preferential loss of short microtubules but by a general shift in the entire distribution. At the lowest concentration measured, the mean length decreased slightly at 8 min, although in this population there were some microtubules longer than in the original population (Fig. 6).

This loss of microtubules following dilution was quite unexpected, as was the continued growth of remaining microtubules. The microtubules could have dropped off, depolymerized, or fallen apart in some catastrophic manner. These hypotheses were tested in an experiment in which regrown centrosomes were diluted to 4.5 μM tubulin, where loss in microtubule number is very rapid. By fixing 15 and 30 s after dilution, it was possible to trap the intermediates in microtubule loss as large numbers of short microtubules. Thus, the microtubule loss following dilution appeared to result from endwise depolymerization—presumably from the distal end, although proximal loss remained a formal possibility. We conclude that at tubulin concentrations below the steady-state level, microtubules on centrosomes exist as two populations, some growing and some shrinking; the fraction in the two populations depends on the free tubulin concentration.

One possible explanation for this behaviour would be structural differences between growing and shrinking microtubules. Conceivably, the centrosome could nucleate microtubules with different stabilities by imposing some structural constraint on the microtubule lattice. The existence of such different microtubule classes could explain the heterogeneous data of Figs 4 and 6. If this were the case, the microtubules nucleated at a low tubulin concentration should be more stable and have a lower critical concentration than those nucleated at higher concentration. We tested this hypothesis by initiating polymerization on centrosomes at a tubulin concentration giving about half the saturating number of microtubules. After microtubules nucleated at this concentration had grown to $\sim 6 \mu\text{m}$, more tubulin was added to reinitiate microtubule assembly, which resulted in a bimodal initial length distribution (Fig. 7). The length distribution on any given centrosome was also bimodal. After dilution, the populations remained bimodal, with equal fractions in each population, despite loss of over half the microtubules (Fig. 7). This is clearly inconsistent with the hypothesis. The peaks even appear to move up slightly in length during this short time, illustrating again that the remaining microtubules continued to grow. It thus seemed that microtubules depolymerize at random following dilution, and furthermore that depolymerization, once initiated, continues to completion. If depolymerizing microtubules had a significant probability of reverting to polymerizing behaviour, the upper and lower peaks would have merged. This is consistent with the data of Fig. 6, which also suggests that the shrinking and growing microtubules are not readily interconvertible. If they were, the mean length of the remaining microtubules would be much lower than found.

Conclusions

The data described here indicate that the centrosome is capable of microtubule nucleation at tubulin concentrations where individual microtubules are unstable and can only exist transiently. A newly nucleated microtubule can grow for a while but will eventually start to depolymerize with a probability which increases with decreasing concentration of tubulin. Once initiated, the microtubule will generally disappear completely, leaving an unoccupied nucleation site. This site can then initiate a new microtubule. Because the centrosome can continuously renucleate, it will always be at the centre of an astral microtubule array, which will thus appear stable. Each individual microtubule is unstable, but a stable amount of polymer is found despite the overall tubulin concentration being below that at which free microtubules are stable.

If such properties are manifest in the cell, they have profound consequences for the structure and dynamics of the microtubule cytoskeleton; they could explain why most of the microtubules in many cell types originate from the centrosome, and why free microtubules, formed for example by removal of depolymerizing drugs, are unstable with respect to centrosomal microtubules¹⁵.

The *in vitro* experiments demonstrate not only that microtubule assembly off the centrosome is kinetically preferred, but also that centrosomes will be the source of microtubules in the cell in the long term. Although differences in the affinity of the ends of microtubules for tubulin could accentuate the stability of centrosomal microtubules near the steady-state concentration¹⁷, the properties observed here could operate well below

the steady-state concentration for either end of a microtubule and in these conditions would predominate.

Further consequences of this dynamic behaviour are considered in the following paper²³, which also considers the effects of the microtubule having two ends. We also propose a model making use of the concept of a cap of GTP-containing subunits at the ends of growing microtubules^{32,33} to account for the coexistence of shrinking and growing microtubules in a single population.

We thank T. Hill for helpful discussion, S. Blose for anti-tubulin antibody, F. McKeon for 5051 anti-centrosome antiserum and Cynthia Cunningham-Hernandez for help in preparing the manuscript. The work was supported by grants from the NIH and the ACS.

Received 27 April; accepted 6 September 1984.

- McIntosh, J. R. *Mod. Cell Biol.* **2**, 115-142 (1983).
- Wheatley, D. N. *The Centriole: A Central Enigma of Cell Biology* (Elsevier, New York 1982).
- Calarco-Gillam, P., Siebert, M., Hubble, R., Mitchison, T. & Kirschner, M. *Cell* **35**, 621-629 (1983).
- Osborn, M. & Weber, K. *Proc. natn. Acad. Sci. U.S.A.* **73**, 867-871 (1976).
- Brinkley, B. R., Fuller, A. M. & Highfield, D. P. in *Cell Motility* (eds Goldman, R., Pollard, T. & Rosenbaum, J.) 867-871 (Cold Spring Harbor Laboratory, New York, 1976).
- Frankel, F. R. *Proc. natn. Acad. Sci. U.S.A.* **73**, 2798-2802 (1976).
- Snyder, J. A. & McIntosh, J. R. *J. Cell Biol.* **67**, 744-760 (1975).
- Brinkley, B. R. *et al. J. Cell Biol.* **90**, 554-562 (1981).
- Gould, R. R. & Borisy, G. G. *J. Cell Biol.* **73**, 601-615 (1977).
- Bergen, L., Kuriyama, R. & Borisy, G. G. *J. Cell Biol.* **84**, 151-159 (1980).
- Kuriyama, R. *J. Cell Sci.* **66**, 277-295 (1984).
- Roobol, A., Havercroft, J. C. & Gull, K. *J. Cell Sci.* **55**, 365-381 (1982).
- Ring, D., Hubble, R., Caput, D. & Kirschner, M. W. in *Microtubules and Microtubule Inhibitors* (eds DeBrabander, M. & DeMey, J.) 297-309 (Elsevier, Amsterdam, 1980).
- Kuriyama, R. & Borisy, G. G. *J. Cell Biol.* **91**, 822-826 (1981).
- DeBrabander, M., Geuens, G., DeMey, J. & Joniau, M. in *Microtubules and Microtubule Inhibitors* (eds DeBrabander, M. & DeMey, J.) 255-268 (Elsevier, Amsterdam, 1980).
- DeBrabander, M., Geuens, G., Nuydens, R., Willebrords, R. & DeMey, J. *Cell Biol. int. Rep.* **5**, 913-920 (1981).
- Kirschner, M. W. *J. Cell Biol.* **86**, 330-334 (1980).
- Hill, T. L. & Kirschner, M. W. *Int. Rev. Cytol.* **84**, 185-234 (1982).
- Spiegelman, B. M., Lopata, M. & Kirschner, M. W. *Cell* **16**, 239-252 (1979).
- Sharp, G. A., Osborn, M. & Weber, K. *J. Cell Sci.* **47**, 1-24 (1981).
- Ring, D., Hubble, R. & Kirschner, M. W. *J. Cell Biol.* **94**, 549-556 (1982).
- Karsenti, E., Newport, J., Hubble, R. & Kirschner, M. W. *J. Cell Biol.* **98**, 1730-1745 (1984).
- Mitchison, T. & Kirschner, M. W. *Nature* **312**, 237-242 (1984).
- Lee, J. C. & Timasheff, S. N. *Biochemistry* **14**, 5183-5187 (1975).
- Oosawa, F. & Asakura, S. *Thermodynamics of the Polymerization of Protein* (Academic, New York, 1975).
- Johnson, K. A. & Borisy, G. G. *J. molec. Biol.* **117**, 1-31 (1974).
- Gaskin, F., Cantor, C. R. & Shelanski, M. L. *J. molec. Biol.* **89**, 737-758 (1974).
- Jameson, L. & Caplow, M. *J. biol. Chem.* **255**, 2284-2292 (1980).
- Karr, T. L., Kristofferson, D. & Purich, D. L. *J. biol. Chem.* **255**, 8567-8572 (1980).
- Cote, R. H. & Borisy, G. G. *J. molec. Biol.* **150**, 577-602 (1981).
- Witman, G. B., Cleveland, D. W., Weingarten, M. D. & Kirschner, M. *Proc. natn. Acad. Sci. U.S.A.* **73**, 4070-4074 (1976).
- Carlier, M.-F. & Pantaloni, D. *Biochemistry* **29**, 1918 (1981).
- Carlier, M.-F., Hill, T. & Chen, Y.-D. *Proc. natn. Acad. Sci. U.S.A.* **81**, 771 (1984).
- Blose, S. H., Metzger, D. I. & Feramisco, J. R. *J. Cell Biol.* **98**, 847-858 (1984).
- Weingarten, M. D., Suter, M. S., Littman, D. R. & Kirschner, M. W. *Biochemistry* **13**, 5529-5537 (1974).
- Terry, B. J. & Purich, D. L. *J. biol. Chem.* **254**, 9469-9476 (1979).
- Bradford, M. *Analyt. Biochem.* **72**, 248 (1976).

Dynamic instability of microtubule growth

Tim Mitchison & Marc Kirschner

Department of Biochemistry and Biophysics, University of California at San Francisco, San Francisco, California 94143, USA

We report here that microtubules in vitro coexist in growing and shrinking populations which interconvert rather infrequently. This dynamic instability is a general property of microtubules and may be fundamental in explaining cellular microtubule organization.

MICROTUBULES are structural filaments in the cytoplasm which are spatially organized and extremely dynamic^{1,2}. Recently, considerable effort has been directed towards understanding what produces and stabilizes specific arrangements of microtubules in cells and by what means microtubules can completely reorganize their spatial distribution. In the accompanying paper³, we suggest that microtubules nucleated by centrosomes can grow transiently at tubulin concentrations below those at which free microtubules are stable, and that nucleated microtubules coexist as shrinking and growing populations which rarely interconvert. This behaviour is clear only when individual microtubules rather than bulk populations are studied. Here we generalize the results from microtubules nucleated by centrosomes to free microtubules. We examine the detailed kinetics of microtubule assembly to try to account for these unusual dynamic properties.

Microtubule dilution

The crucial experiment demonstrating unusual dynamics in the preceding paper was that in which the microtubule number and length distributions were measured after centrosomes were regrown initially at a high tubulin concentration, then diluted³. The conclusion we drew from that experiment was that some microtubules continued to grow at the same time as others were lost by depolymerization from their distal ends. It seemed likely that this was a general property of microtubules. We therefore

describe here a similar experiment with free microtubules (Fig. 1). Microtubules were first made by spontaneous polymerization in an assembly-promoting buffer. These were then used as seeds and diluted extensively into a tubulin solution well above the steady-state concentration and allowed to elongate for 4 min. The seeds, which were initially 1-2 μm long, elongated to form a sharp distribution with a mean length of 18.3 μm (Fig. 1b). This actively growing population was either fixed immediately or first diluted with warm buffer to just above (15 μM) or below (7.5 μM) the steady-state concentration. To assess both the length and number concentration, fixed microtubules were quantitatively sedimented onto grids for electron microscopy. This procedure gave a highly reproducible number concentration (Fig. 1) and when a known polymer mass was used it gave the expected mean length (see, for example, Fig. 4).

The result of this experiment was very similar to that found using centrosome nucleated microtubules³, that is, below the steady-state concentration microtubules were found to both grow and shrink (Fig. 1). Above the steady-state concentration, the number of microtubules remained approximately constant (Fig. 1a) and their length increased from 18.3 to 40.2 μm in 10 min, retaining a fairly sharp distribution (Fig. 1d). Below the steady-state concentration, however, the number concentration decreased with time (Fig. 1a), but the mean length still increased from 18.3 to 21.5 μm in 10 min (Fig. 1c). The rate of microtubule

disappearance was similar to that of centrosomal microtubules at the same tubulin concentration, suggesting that a free minus end does not markedly decrease microtubule stability. This experiment demonstrates the coexistence of shrinking and growing populations with free microtubules and the transient growth of microtubules below the steady-state concentration, confirming that the centrosome data were demonstrating a general property of microtubules. The net polymer mass is clearly decreasing at the lower concentration, as the increase in mean length is more than offset by the decrease in number. At the higher concentration, the net polymer mass increases, confirming that the steady-state concentration lies between these concentrations (Fig. 1a). Since no renucleation occurs in this experiment, polymer mass will eventually decrease to zero at the lower concentration. Centrosomes present at the same concentration would, however, continuously renucleate microtubules³.

Microtubule polymerization rates

In order to understand the dynamic processes occurring in the dilution experiment, we sought to determine the quantitative relationship between tubulin concentration and microtubule polymerization and depolymerization rates. A method giving data for plus and minus ends independently was chosen⁴. To determine polymerization rates, flagellar axonemes were mixed with prewarmed tubulin solutions, then fixed at various time intervals and the length of the microtubules polymerized onto each end of the axonemes was determined (the ends of the axoneme are distinguishable in the electron microscope⁵). To determine depolymerization rates, axonemes were first regrown with microtubules, then diluted (Fig. 2). Immunofluorescent visualization was used for higher tubulin concentrations (Fig. 2a) and electron microscopy for lower concentrations (Fig. 2b, c), with equivalent results. For the higher tubulin concentrations, the length increase is linear up to $\sim 20 \mu\text{m}$, at which length shear-induced breakage becomes difficult to avoid (Fig. 2a). At the lowest concentrations the mean length rapidly plateaued and only initial rates were used (Fig. 2b). The linear decrease of length with time (Fig. 2c) is consistent with endwise depolymerization.

Figure 3 plots the rate measurements as a function of tubulin concentration. The points on the positive portion of the data, which we refer to as the growing phase, show a linear relationship between growth rate and concentration. This is interpreted in terms of the simple first-order rate equation shown in the legend to Table 1. The values of k_T for the plus and minus end are quite similar to those derived for microtubule protein^{4,6}. As in those studies, the fast-growing end corresponds to the end of the axoneme distal to the cell.

The data cannot be used to derive the x intercepts very accurately because they are so close to zero, and small changes in the slope produce large changes in their values. The data do not support the assertion that the intercepts for the plus and minus ends are significantly different. Thus, the growing phase

Table 1 Derived rate constants for the kinetics of microtubule assembly

	Plus end	Minus end	Units
Slope	0.135	0.042	$\mu\text{m } \mu\text{M}^{-1} \text{ min}^{-1}$
k_T	3.82	1.22	$\times 10^6 \text{ M}^{-1} \text{ s}^{-1}$
x intercept	0.1	0.9	μM
y intercept	-0.013	-0.039	$\mu\text{m min}^{-1}$
k_T'	0.37	1.1	s^{-1}
Depolymerization rate	-12.0	-7.5	$\mu\text{m min}^{-1}$
k_D'	340	212	s^{-1}

We assume 1,700 tubulin dimers per μm microtubule. The growth rate (J_T) is defined as $k_T c - k_T' d$, where c is the tubulin monomer concentration, k_T the second-order on-rate constant and k_T' the extrapolated off-rate constant during polymerization, calculated from y intercept of Fig. 3.

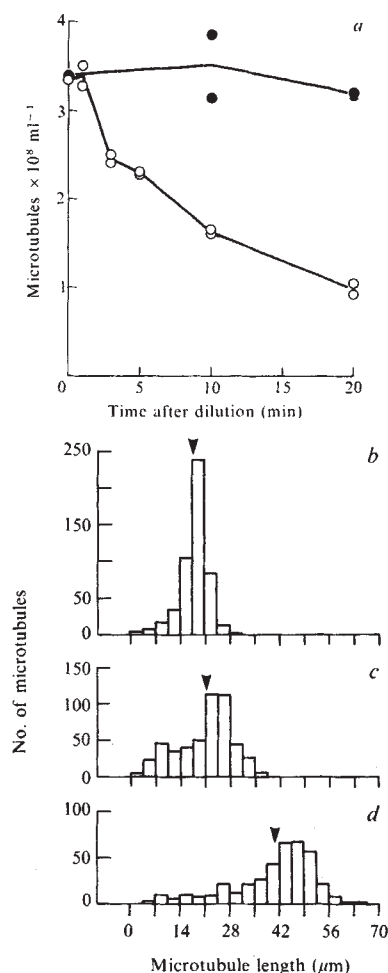


Fig. 1 Dilution after seeded assembly. *a*, For number concentration, the number of microtubule ends per field was averaged for 16 fields, divided by two and scaled to final number concentration using the geometry of the rotor and the dilution factor. Duplicate grids were made at each time point. Final tubulin concentration: closed circles $15 \mu\text{M}$, open circles $7.5 \mu\text{M}$. Microtubule lengths were determined by digitizing as in Fig. 6 of ref. 3. *b-d*, Length histograms. *b*, Fixed before dilution (500 measured, mean $18.3 \mu\text{m}$, s.d. = $4.0 \mu\text{m}$); *c*, 10 min at $7.5 \mu\text{M}$ (500 measured, mean $21.5 \mu\text{m}$, s.d. = $8.6 \mu\text{m}$). The mean is significantly increased ($P < 0.0001$) with respect to the starting population. *d*, 10 min at $15 \mu\text{M}$ (250 measured, mean $40.2 \mu\text{m}$, s.d. = $12.0 \mu\text{m}$). The arrow points to the mean of each distribution.

Methods: Preparation of tubulin, measurement of protein concentration, buffers used and generation of microtubule seeds were as described in Fig. 5 of ref. 3. 1/100 volume of seeds was added to prewarmed tubulin, $25 \mu\text{M}$ in PB. After 4 min at 37°C , $10\text{-}\mu\text{l}$ aliquots of the growing microtubules were added to $100 \mu\text{l}$ of 1% glutaraldehyde in PB' (PB without GTP) at 26°C or 1 ml of prewarmed tubulin at $15 \mu\text{M}$ (above steady-state concentration) or $7.5 \mu\text{M}$ (below steady-state concentration) in PB, and gently mixed. Incubation was continued at 37°C , and aliquots were fixed at the indicated time intervals. After 3 min in the fixative, the microtubules were diluted with cold PB', and sedimented onto polylysine-coated 150-mesh Parlodion grids in the airfuge EM 90 rotor (Beckman) at $90,000g$ for 15 min. Grids were previously polylysine-coated by immersion in $1 \mu\text{g ml}^{-1}$ polylysine for 10 min, then dried by aspiration. Grids were dried and shadowed as described in Fig. 4 of ref. 3. Random fields were photographed at a final magnification of $\times 1,500$.

of microtubules behaves like simple subunit addition to a polymer in equilibrium with its subunits, having the same extremely small critical concentrations for the two ends⁷. In any case, the x intercepts for both ends are much lower than the steady-state concentration measured on the same tubulin preparation³, which was $14 \mu\text{M}$. Thus axonemes, like centrosomes, can nucleate

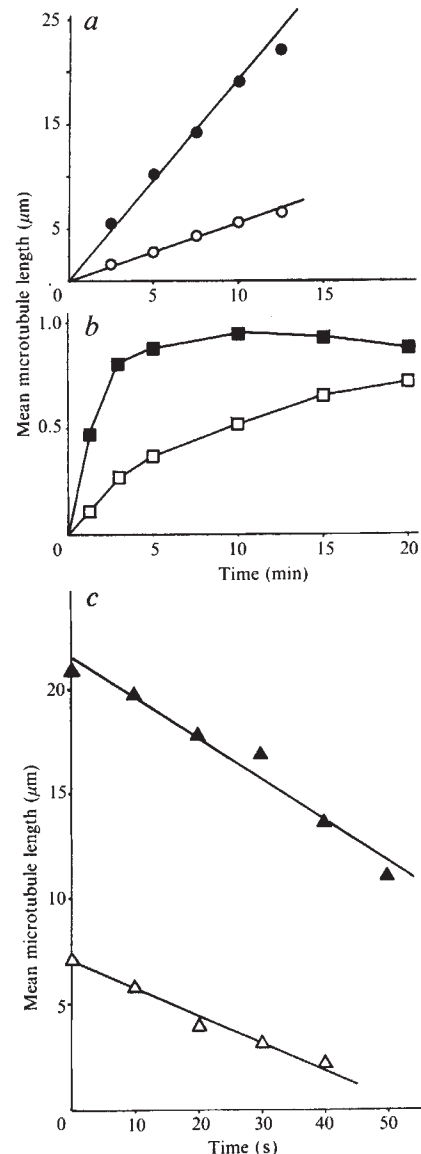
Fig. 2 Microtubule growth rate off axonemes. *a*, Polymerization at 15 μM tubulin (immunofluorescence). *b*, Polymerization at 3 μM tubulin (electron microscopy). *c*, Depolymerization at 2.5 μM tubulin (electron microscopy).

Methods: Axonemes were prepared by washing *Tetrahymena* cells in fresh growth medium, followed by 10 mM PIPES, 1 mM MgCl_2 , 1 mM CaCl_2 , pH 7.2 with KOH. Cells were resuspended at 25 $^\circ\text{C}$ in the same buffer and dibucaine was added to 0.5 mM²⁴. After 5 min the cells shed their axonemes and rounded up. Cell bodies were removed by centrifuging twice at 2,000g for 5 min. The supernatant was made 5 mM in EDTA and 0.5% Triton X-100. Axonemes were pelleted at 25,000g for 20 min at 4 $^\circ\text{C}$ and resuspended in 50% (v/v) glycerol, 5 mM PIPES, 0.5 mM EDTA, 1 mM β -mercaptoethanol, pH 7 with KOH at a concentration of $\sim 10^{10} \text{ ml}^{-1}$. They could be stored for months at -20 $^\circ\text{C}$ in this buffer. Solutions of tubulin in PB were prewarmed to 37 $^\circ\text{C}$ for 2 min and axonemes were added to a final concentration of 10^7 ml^{-1} . The mixture was incubated at 37 $^\circ\text{C}$ and aliquots were fixed as in Fig. 1 at five or six time points. The time course varied from 20 min at the lower to 5 min at the higher concentrations. Fixed regrown axonemes were diluted and sedimented onto polylysine-coated grids in the airfuge as in Fig. 1, or onto polylysine-coated glass coverslips (Fig. 1b of ref. 3) except that the spin was for 30 min. Regrown axonemes were then visualized by rotary shadowing and immunofluorescence respectively. For depolymerization experiments, axonemes were pregrown to $\sim 20 \mu\text{m}$ on the plus and $\sim 7.5 \mu\text{m}$ on the minus ends. They were then diluted and fixed at 10-s intervals and analysed by electron microscopy. Random axonemes were photographed at $\times 1,500$ or $\times 3,000$ in the electron microscope, or at $\times 250$ in the light microscope and digitized directly from the negatives. At least 100 microtubules from each end were measured at each time point and the time courses were plotted. Rate of microtubule growth was determined by linear regression to the linear portion of the time course, which generally included all points with mean length less than 20 μm , with typical correlation coefficients of >0.95 .

microtubules well below the steady-state concentration, and this can occur on both the plus and minus ends. This also confirms the data of Fig. 1, where microtubules with both plus and minus ends free continued to grow well below the steady-state concentration.

The shrinking phase of Fig. 3 confirms and extends earlier data⁸ showing a large break in the bulk polymerization curve (measured by turbidity) where it crossed the x axis. For both the plus and minus ends, the observed depolymerization rate is 2–3 orders of magnitude greater than the extrapolated y intercept (k'_T) from the growing phase of the graph. Thus the off rates during depolymerization, which we call k'_D , are much larger than the off rates during polymerization, k'_T . We interpret the difference as being due to the existence of a GTP cap during polymerization, and hence the nomenclature of k'_T and k'_D where k'_T is the off rate of GTP tubulin from a GTP lattice, and k'_D is the off rate of GDP tubulin from a GDP lattice, in agreement with earlier studies⁸.

It is important to distinguish the kind of data obtained by bulk measurements⁸ from the measurement here of individual microtubules. The growing phase data of Fig. 2 is not an aggregate growth rate of both growing and shrinking microtubules because the axoneme template could not depolymerize and thus we only observed net growth. At most of the tubulin concentrations used, microtubules grew continuously during the time course, indicating that the transit of a growing microtubule into the shrinking phase is a relatively rare event. Only at 3 and 4 μM tubulin did phase transitions during the time course become significant. At these concentrations, microtubules depolymerized after transient growth, leading to plateauing of average length and a heterogeneous length distribution. Similarly, during the depolymerization experiments, the tubulin concentrations were sufficiently low that essentially all the microtubules shrank continuously. Thus the two arms of the plot in Fig. 2 cannot be connected. A single shrinking microtubule can transit from the growing phase to the shrinking phase over a wide range of concentrations (Fig. 1; ref. 3).



Steady-state dynamics

The dynamics at steady state are unclear, because at 14 μM tubulin (where the net polymer mass should be constant) the plus and minus ends should be growing at 1.88 and 0.55 $\mu\text{m min}^{-1}$ respectively, or a total of 2.43 $\mu\text{m min}^{-1}$ for free microtubules. We examined the nature of the steady state by measuring microtubule length and number concentration (Fig. 4). Microtubule polymerization was initiated by seeded assembly and followed by turbidity. After assembly to near steady state the microtubules were sheared. Following a transient decrease in turbidity, the microtubules rapidly attain a plateau in turbidity, corresponding to steady-state assembly (Fig. 4a). Sampling during the plateau of turbidity indicated that there is a steady increase in mean microtubule length, while the number concentration decreases steadily. Polymer concentration determined as the product of average length and number is constant (within an experimental error of 10%) and the mean polymer concentration (30 μM) is in good agreement with that expected from the bulk assembly data in our previous paper (Fig. 5; ref. 3).

The increase of mean length with time demonstrates that most microtubules are indeed growing at the steady-state concentration. Monomer for this growth must be supplied by the depolymerization of shrinking microtubules, and in the absence of renucleation this leads to the observed decline in number concentration. Thus the steady state can be interpreted as being due to the coexistence of two phases, with the majority of microtubules growing slowly balanced by the minority shrinking

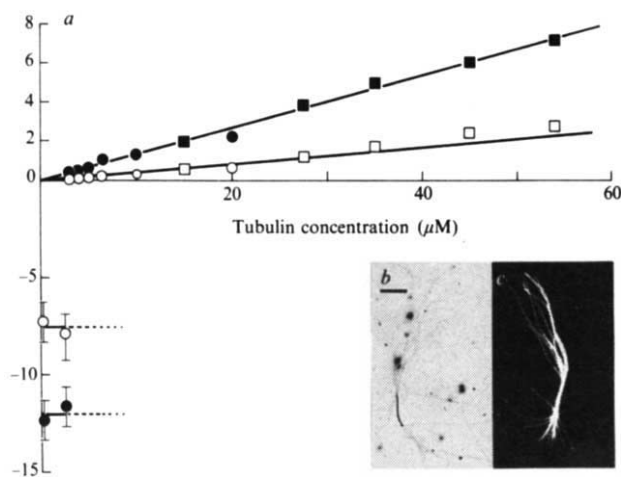


Fig. 3 *a*, The growth rate of microtubules off axonemes obtained from data of the type in Fig. 2 plotted as a function of tubulin concentration. The closed symbols show data for the plus and the open for the minus ends. Data points plotted as squares were determined by immunofluorescence, and points plotted as circles by electron microscopy and shadowing. The line is a least-squares fit which minimizes relative deviations, that is $(di/yi)^2$ rather than $(di)^2$. All points have similar relative errors thus points at the lower tubulin concentration have smaller absolute errors, so minimizing relative deviation gives the most physically reasonable fit to the data. *b*, Typical regrown axoneme visualized by rotary shadowing and electron microscopy. Scale bar, 4.5 μm. *c*, Typical regrown axoneme visualized by immunofluorescence and printed to the same scale as *b*.

rapidly. The observed length fluctuations are much too large to be explained by random fluctuations of an equilibrium polymer. Using reasonable rate constants, a polymer 15 μm long would take a year to fluctuate to zero by fluctuations at equilibrium^{9,10} and the addition of treadmilling would not give the observed rates¹¹. End-to-end ligation of microtubules is also unlikely to explain the results, as experiments with pulses of biotin-labelled tubulin have shown only end addition as the mechanism of elongation (manuscript in preparation). The proportion in the two phases can be estimated from Fig. 4, because net growth must balance net disassembly at steady state. Unfortunately, the existence of different kinetics at the two ends complicates the analysis, but if most of the net growth occurs on plus ends at 1.9 μm min⁻¹ (the growth rate at 14 μM tubulin dimer) and disassembly occurs off both ends at an average of 9.7 μm min⁻¹, then there will be an average 5.2 growing microtubules for every shrinking one.

If transitions between growing and shrinking phases are very frequent, then little net microtubule elongation (or number loss) would be seen. The observed growth rate (Fig. 4) suggests that such transitions must be rare. A crude way to estimate the transition probabilities is to consider that for microtubules ~20 μm long, the half life for loss in number is ~20 min at the plateau in turbidity. Using the kinetic constants described above, such a microtubule would take ~1.7 min to depolymerize from the plus end, or ~2.7 from the minus end. If ~20% of the microtubules are depolymerizing at any one time, the half life should be ~10 min if they always depolymerize to completion once shrinking is initiated. The observed half life of 20 min may reflect that, on average, half the microtubules depolymerize to completion and half are recapped somewhere during the 20 μm of depolymerization, which represents 34,000 dissociation events. However, determination of the transition rates will require further experiments and detailed modelling.

We favour a model which explains the difference between growing and shrinking microtubules in terms of differences at their ends, in that growing microtubules have GTP-liganded caps whereas shrinking microtubules do not⁸. One observation which supports the idea that growing microtubules are stabilized by their GTP caps is shown in the absorbance traces in Fig. 4.

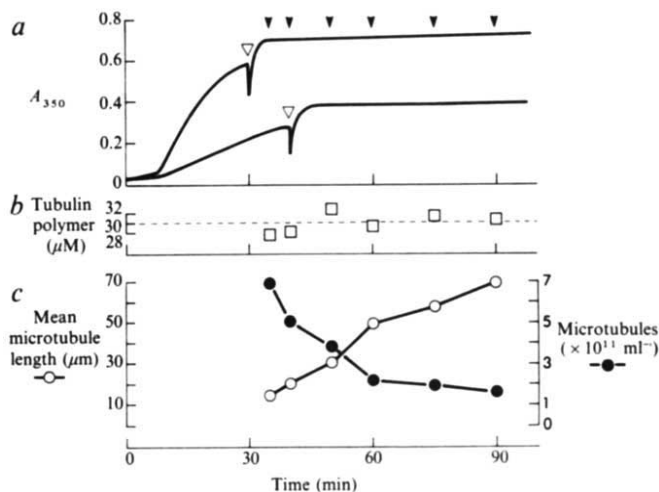


Fig. 4 Length redistribution at steady state. *a*, Turbidity record: upper trace, total tubulin = 59 μM; lower trace, total tubulin = 32 μM. Microtubules sampled at times shown by arrows. *b*, The product of mean length and number concentration, expressed as μM polymer, assuming 1,700 subunits μm⁻¹. *c*, Number concentration and mean length at the time point denoted by the arrows in *a*, upper trace. The same time scale is used for *a*, *b* and *c*; the points in *c* refer to the samples denoted by closed arrows in *a*. *d*, *e*, *f*, Part of typical immunofluorescence fields. Arrow = 30 μm. Time after shear: *d*, 5 min; *e*, 15 min; *f*, 60 min.

Methods: Microtubule seeds were prepared as in Fig. 5 of ref. 1. Tubulin in PB was added to a prewarmed cuvette in a Cary spectrophotometer and incubated throughout at 37°C. At 8 min, 1/100 volume of seeds was added and the solution mixed. At the time indicated by the open arrow, the growing microtubules were sheared by passing the solution twice through a 22 gauge 1½-inch needle, using a 1-ml syringe which had been prewarmed to 37°C. From the solution in the upper trace, aliquots were removed for fixation at the times indicated by closed arrows. Aliquots (20 μl) were added to 2 ml of 1% glutaraldehyde in PB' at 26°C. After 3 min, the fixed microtubules were further diluted using cold PB'. Wide-mouthed pipette tips and very gentle mixing were used to minimize shear. 100-μl aliquots of diluted microtubules were spun onto 4.5-mm² polylysine-coated glass coverslips using the airfuge EM 90 rotor at 90,000g for 15 min. The microtubules were visualized by anti-tubulin immunofluorescence without methanol post-fixation. Microtubules were photographed at ×250 and ×100, and digitized for fluorescent axonemes as in Fig. 2. For the length data, 500 microtubules were measured at each time point. The higher magnification was used for the first three points, and the lower for the last three. For the number data, the number of microtubule ends per field at ×250 was averaged over 38 fields, divided by two and scaled using the magnification, the geometry of the rotor and the dilution factor.

When the seeded assembly mixture is sheared during the growth phase, there is a large transient decrease in the A_{350} , indicating considerable microtubule depolymerization. The extent of this drop, which can be up to one-third of the polymer present, depends on the extent of shearing. Merely drawing the solution into a Pasteur pipette has little effect on the A_{350} . Such a shear-induced depolymerization is not expected from simple theories of polymer polymerization⁹. We interpret this transient depolymerization as being due to the breakage of microtubules, which exposes GDP-liganded subunits at the new ends. These are unstable and start to depolymerize rapidly, a process which continues until new GTP caps can be established.

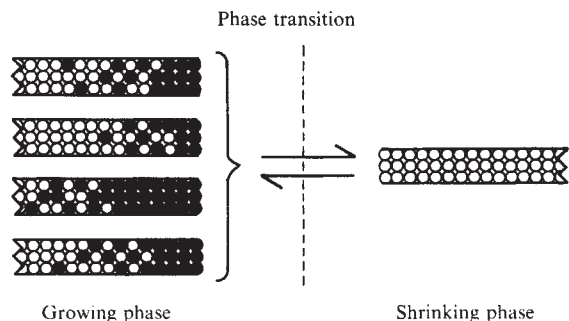


Fig. 5 Model for microtubule ends exchanging subunits with soluble dimer. The open circles represent tubulin liganded with GDP, and the closed circles represent GTP-tubulin. The growing phase of microtubules is stabilized by a GTP cap. This cap is of fluctuating length, but the average length increases with free tubulin concentration. The shrinking phase has lost this cap. Interconversion between the two phases is relatively rare.

Microtubule assembly

A new model for microtubule assembly arising from the data presented here and other data^{3,8,12} is presented in Fig. 5. The essence of the model is that two distinct phases of microtubules exist which are distinguished by the presence or absence of a GTP-liganded cap, and that the two phases interconvert infrequently. The ends without a cap are unstable at all monomer concentrations tested, that is, GTP subunits will not give net addition to a GDP lattice in the concentration range studied. The GDP lattice thus depolymerizes rapidly, with off rates >100 times faster than the y intercept extrapolated from the growing phase kinetics (Table 1). The GTP-capped lattice, however, is stabilized, because the off rate of GTP subunits is slower by 2–3 orders of magnitude, so net addition of new GTP subunits occurs down to a very low tubulin concentration. However, the cap only exists because the hydrolysis rate lags slightly behind the polymerization reaction¹³. GTP hydrolysis is a first-order reaction initiated by the conformational change undergone by the subunit once incorporated into the polymer lattice. Thus, once a GTP subunit is incorporated, there is a fixed probability per unit time of hydrolysis occurring. Addition of GTP subunits, however, is a second-order reaction whose rate depends on monomer concentration. At high tubulin concentration there will always be a significant cap of GTP subunits at the growing tip of the polymer. In these circumstances nearly all microtubules will be in the growing phase. However, as the monomer concentration is decreased, the average cap size will decrease, statistical fluctuations in cap size will become important and the probability of GDP subunits becoming exposed on or near the polymer end increases. When this happens, the end subunits are rapidly lost into solution because their affinity for the GDP lattice is low. Once rapid depolymerization is initiated, it continuously exposes fresh GDP subunits at the polymer end and continues until the polymer disappears or a rare event recaps the microtubule and growth restarts. The events that cause a shrinking microtubule to become recapped are not clear. Depolymerization may terminate when encountering remaining GTP in the lattice, and the shrinking end could then become recapped by further addition of GTP-tubulin. The shrinking GDP end may also be capped by binding of GTP tubulin to the GDP end or by direct exchange of GDP for GTP on the terminal subunit.

It is clear, however, that phase transitions in either direction are rare compared with subunit addition in the growing phase or loss in the shrinking phase. Both phase transition probabilities should be quite sensitive to monomer concentration, and perhaps also to the free nucleotide concentrations. An explicit kinetic model incorporating these features has been developed recently¹⁴. Using plausible rate constants and Monte Carlo calculations, the essential features of persistence of growth and

shrinking (with rare interconversions) have been demonstrated and a general model constructed¹⁵.

An important question concerns the effect of other components of microtubules, such as microtubule-associated proteins (MAPs), on this dynamic behaviour. It has been suggested⁸ that MAP-containing microtubules also have a GTP cap, as they also have a large discontinuity between growing and shrinking kinetics. Some experiments have demonstrated appreciable length fluctuations at steady state¹¹. As MAPs bind to the microtubule lattice and stabilize it, they could slow the loss of GDP subunits and increase the probability that a shrinking microtubule would transit into the growing phase. This, in turn, would lower the steady-state concentration. The greater number of transitions at steady state induced by MAPs should tend to damp out the observed length fluctuations. Another factor which would tend to obscure length redistribution at steady state by offsetting the loss of microtubule number is some renucleation of microtubules, which may be strongly promoted by MAPs.

These considerations question the interpretation of isotope uptake experiments by microtubules at steady state as the result of a flux of subunits through the polymer or treadmilling^{16,17}. The growing microtubule may behave instead like a simple equilibrium polymer, and extensive isotope uptake may occur through steady-state length redistribution. For microtubules with MAPs, length redistributions (and isotope uptake) would be expected to be less extensive, and in fact the measured isotope uptake is very small for tubulin-containing MAPs¹⁶ compared with purer preparations of tubulin¹⁸.

Such issues are important because of the question of the relevance of *in vitro* dynamic data with pure tubulin to the living cell. The fast depolymerization rate of microtubules as compared to the extrapolated values from the growing phase holds under various *in vitro* circumstances^{8,11,19,20}, and is likely to be true also *in vivo*. This may mean that the primary reason the cell has evolved GTP hydrolysis by tubulin is to have a polymer with built-in instability. Such a polymer could grow rapidly and be stabilized by end interactions, but it could also shrink rapidly, regulated by only small changes in conditions. The ability of microtubules to shrink rapidly may be of importance in reorienting the interphase microtubule array during locomotion or during morphological changes and in the extensive microtubule rearrangements of mitosis². For example, using the off rates for the growing phase, a 20 μ m microtubule would take more than 6 h to depolymerize at zero tubulin concentration, but using the off rates for the shrinking phase it would take about 1 min. The importance of the GTP cap in stabilizing growing microtubules raises the possibility that capping proteins may exist in cells and could be required for the long-term stabilization of a microtubule in a cell. Thus the centrosomes may be continually nucleating new microtubules, and only those ends which become capped or become stabilized in some other way such as by inhibition GTP hydrolysis may persist for long periods. A particular example of this could be the capture and capping of centrosomal microtubules by the kinetochore in prometaphase^{21,22}.

Various techniques have been used to study the dynamics of microtubules in living cells, and recently the powerful method of introducing fluorescently labelled tubulin molecules into living cells has been exploited. Recent results show that the mitotic spindle is much more dynamic than expected²³ and complete exchange of spindle microtubule and soluble subunits occurs within seconds. Interphase microtubules are considerably more stable but still exchange with soluble subunits within minutes. The dynamic behaviour induced in microtubules by the presence of fluctuating GTP caps provides the only satisfactory *in vitro* explanation for the rapid exchange of spindle microtubules.

We thank T. Hill, M.-P. Carlier and D. Pantaloni for helpful discussion, H. Martinez for the linear regression used in Fig. 2, and Cynthia Cunningham-Hernandez for help in preparing the manuscript. This work was supported by grants from the NIH and the ACS.

Received 27 April; accepted 6 September 1984.

1. Roberts, K. & Hyams, J. *Microtubules* (Academic, London, 1977).
2. McIntosh, J. R. *Mod. Cell Biol.* **2**, 115-142 (1983).
3. Mitchison, T. J. & Kirschner, M. *Nature* **312**, 232-237 (1984).
4. Bergen, L. G. & Borisy, G. G. *J. Cell Biol.* **84**, 141-150 (1980).
5. Allen, C. & Borisy, G. G. *J. molec. Biol.* **90**, 381-402 (1974).
6. Summers, K. & Kirschner, M. W. *J. Cell Biol.* **83**, 205-217 (1979).
7. Hill, T. L. & Kirschner, M. W. *Int. Rev. Cytol.* **84**, 185-234.
8. Carlier, M.-F., Hill, T. L. & Chen, Y.-D. *Proc. natn. Acad. Sci. U.S.A.* **81**, 771-775 (1984).
9. Oosawa, F. & Asakura, S. *Thermodynamics of the Polymerization of Protein* (Academic, New York, 1975).
10. Carlier, M.-F., Pantaloni, D. & Korn, E. D. *J. biol. Chem.* (in the press).
11. Kristofferson, D. & Purich, D. L. *Archs Biochem. Biophys.* **211**, 222-226 (1981).

12. Hill, T. L. & Carlier, M.-F. *Proc. natn. Acad. Sci. U.S.A.* **80**, 7234-7238 (1983).
13. Carlier, M.-F. & Pantaloni, D. *Biochemistry* **20**, 1918-1924 (1981).
14. Hill, T. L. & Chen, Y.-D. *Proc. natn. Acad. Sci. U.S.A.* (in the press).
15. Hill, T. *Proc. natn. Acad. Sci. U.S.A.* (in the press).
16. Margolis, R. L. & Wilson, L. *Cell* **13**, 1-8 (1978).
17. Farrell, K. W. & Jordan, M. A. *J. biol. Chem.* **257**, 3131-3138 (1982).
18. Cote, R. H. & Borisy, G. G. *J. molec. Biol.* **150**, 577-602 (1981).
19. Jameson, L. & Caplow, M. *J. biol. Chem.* **255**, 2284-2292 (1980).
20. Karr, T. L., Kristofferson, D. & Purich, D. L. *J. biol. Chem.* **255**, 8560-8565 (1980).
21. Pickett-Heaps, J., Tippit, D. H. & Porter, K. R. *Cell* **29**, 729-744 (1982).
22. Mitchison, T. J. & Kirschner, M. W. in *Molecular Biology of the Cytoskeleton* (eds Cleveland, D. W., Murphy, D. & Borisy, G. G.) (Cold Spring Harbor Laboratory, New York, in the press).
23. McIntosh, J. & Salmon, E. D. *J. Cell Biol.* (in the press).
24. Thompson, G. A., Baugh, C. C. & Walker, L. F. *J. Cell Biol.* **61**, 253-257 (1974).

Polyoma virus DNA replication requires an enhancer

Jean de Villiers^{*§}, Walter Schaffner^{*}, Chiara Tyndall^{†§},
Steven Lupton^{‡§} & Robert Kamen[‡]

^{*} Institut für Molekularbiologie II der Universität Zürich, Honggerberg CH 8093, Switzerland

[†] ICRF Laboratories, Lincoln's Inn Fields, London WC2A 3PX, UK

[‡] Genetics Institute, 225 Longwood Avenue, Boston, Massachusetts 02115, USA

Sequences which activate polyoma virus DNA replication are located within a region that also includes the transcriptional enhancer. We demonstrate a cis involvement of enhancers in DNA replication by showing that this region can be replaced by other enhancers, in a position- and orientation-independent manner, and that an immunoglobulin gene enhancer confers tissue-specific replicatory ability.

UPSTREAM sequences essential for the expression of the early genes of the DNA tumour viruses, simian virus 40 (SV40)^{1,2} and polyoma virus³, are the archetypes of gene control elements now commonly called transcriptional enhancers. They are *cis*-acting DNA sequences which, when isolated and linked to a variety of genes, dramatically increase expression without increasing template copy number⁴⁻⁶.

Enhancers have been identified within the genomes of SV40^{4,6}, polyoma virus⁵, bovine papillomavirus⁷, within the long terminal repeats of RNA tumour viruses⁸⁻¹¹ and also within other DNA virus genomes¹²⁻¹⁴. Recently, similar control elements have been shown to be involved in the control of expression of the immunoglobulin genes¹⁵⁻¹⁹ and putative enhancers have been identified in the controlling regions of the human and rat insulin genes and the rat chymotrypsin gene²⁰.

Enhancers share a number of properties which together comprise an operational definition. They function over extremely large distances (in some cases, more than several kilobase pairs) and usually in either orientation⁴⁻⁶. Transcription is potentiated at both natural and pseudo-promoters and, although activation of transcription from both proximal and distal promoters can be detected, proximal promoters appear to be preferentially affected^{21,22}. Enhancers show specificity, ranging from the cell-type preference exhibited by papovavirus and retrovirus enhancers^{8,23} to the extreme specificity of the immunoglobulin enhancers for cells of lymphoid origin¹⁵⁻¹⁹. (For recent reviews see refs 24 and 25.)

The polyoma virus enhancer region has a unique function not yet attributed to other viral or cellular enhancers. It includes *cis* sequence elements essential for viral DNA replication^{3,26-28}. Only DNA molecules that have both the polyoma virus origin of replication²⁷⁻³¹ and adjacent (but non-overlapping) sequences

within the enhancer region can replicate in mouse COP-5 cells³, despite the fact that these cells constitutively express the viral large-tumour (T) antigen. (This viral early gene product is essential for polyoma virus DNA replication.) We have asked whether the activation of DNA replication is an inherent property of enhancers *per se*, or an independent function of other element(s) within this region of the polyoma virus genome. We have therefore constructed hybrid viral genomes in which the polyoma virus enhancer is replaced by other DNA fragments, to identify heterologous sequences which could substitute for the replicatory activation component.

Replacement of polyoma virus enhancer

We started with a plasmid recombinant containing the entire 5.3-kilobase (kb) genome of A2 strain polyoma virus (pPy) and deleted, from the regulatory region, the 244-base pair (bp) *Bcl*I to *Pvu*II DNA segment which carries both the transcriptional enhancer⁵ and replicatory activator functions^{3,27,28} (pPyPBΔ; Fig. 1). Although the origin of DNA replication (ORI) is still present, pPyPBΔ neither expresses its early genes nor replicates even if T antigen is provided (Fig. 1). We then inserted various fragments of DNA from the analogous region of (SV40) at the point of the deletion, because the regulatory region of this closely related virus has distinctive repeated sequence elements (Fig. 1) which are well characterized and have apparently separable functions. We determined the ability of the hybrid DNAs to replicate, independently of the requirement for synthesis of the polyoma virus early gene product T antigen, by assaying DNA replication after transfection of COP-5 cells. These polyoma virus-transformed mouse fibroblast cells constitutively produce the viral early proteins and can replicate plasmids containing only the *cis*-acting elements necessary for viral DNA synthesis³. We also assayed DNA replication in mouse 3T6 cells, in which DNA replication is dependent on the ability of the hybrids to express sufficient T antigen and therefore requires both transcriptional enhancer and replicatory activator functions. DNA replication was measured by digestion of extracted DNA with

§ Present addresses: Genetics Institute, 225 Longwood Avenue, Boston, Massachusetts 02115, USA (J.deV.); CSIRO Genetics Research Laboratories, North Ryde, New South Wales, Australia (C.T.); Department of Microbiology, State University of New York, Stony Brook 11794, USA (S.L.).

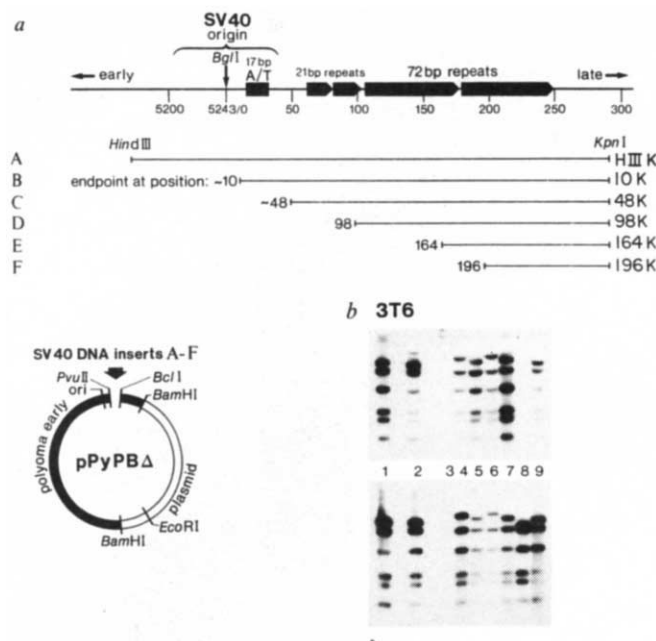


Fig. 1 DNA replication assays in 3T6 and COP-5 cells. **a**, Schematic representation of SV40 DNA fragments inserted into pPyPBΔ, the polyoma virus-plasmid recombinant which has the entire enhancer region deleted (from the *PvuII* site at 5,265 to the *BclI* site at 5,021; numbering of Tooze³¹) and replaced by a synthetic *XhoI* linker fragment. At the top is shown an expansion of the SV40 regulatory region. Fragment A is from the *HindIII* site at 5,171 to the *KpnI* site at 294 (numbering of Tooze³¹), and spans the origin of DNA replication (ORI). It includes the early promoter, and the 21-bp and 72-bp repeats. Below fragment A are fragments B-F, with end points indicated (provided by L. Olson and J. Banerji), which were derived by *Bal31* exonuclease⁵¹ digestion of *HindIII*-cleaved molecules. **b**, Autoradiograph showing DNA replication in 3T6 and COP-5 cells: lane 1, Py (wild type A2 DNA); lane 2, PyPBX (polyoma virus DNA with *XhoI* linkers inserted at *PvuII* and *BclI* sites, spanning the still present enhancer region); lane 3, PyPBΔ; lane 4, PyPBHIIK+ (PyPBΔ with fragment A insert); lane 5, PyPB10K- (PyPBΔ with fragment B insert); lane 6, PyPB98K(2×)- (PyPBΔ with dimer fragment D insert); lane 7, PyPB164K(2×)+ (PyPBΔ with dimer fragment E insert); lane 8, PyPB196K+ (PyPBΔ with fragment F insert); lane 9, PyPB196K(2×)+ (PyPBΔ with tandem dimer insert of fragment F). (+ Refers to the orientation of the insert, which is as shown in the figure; - refers to the opposite orientation.)

Methods: The polyoma virus hybrid recombinants were constructed by inserting the indicated fragments via synthetic *XhoI* linkers, using standard techniques⁵² in accordance with the NIH Guidelines for Recombinant DNA Research. Mouse 3T6 and COP-5 cells were cultured in the presence of 10% fetal calf serum and transfected with these recombinant DNAs, using DEAE-dextran, as described previously⁵⁰. In this experiment the DNAs were excised from the plasmid vector and recircularized by ligation at low concentration before transfection. Transfected cells were lysed after 2.5 days and replication of the DNA was assessed by acquired sensitivity to *MboI* digestion^{3,53}. Briefly, Hirt supernatants⁵⁴ were prepared, extracted once with phenol, digested with RNase and *MboI* endonuclease (which cannot cut the methylated input DNA), separated by agarose gel electrophoresis, blotted to nitrocellulose filters and hybridized to a cloned viral DNA probe, labelled with ³²P by nick-translation. The specific *MboI* cleavage products seen on the autoradiographs indicate replication of the input DNA.

the restriction endonuclease *MboI*. This enzyme cannot cleave the input⁵² plasmid-propagated DNA but cuts unmethylated progeny molecules generated by replication in animal cells (Figs 1-3; Table 1).

Substitution of the polyoma virus enhancer region by the entire regulatory region of SV40 (*HindIII* to *KpnI* fragment A, Fig. 1) including the early promoter, the ORI, the 21-bp repeats and the 72-bp repeat, restored both DNA replication and early gene expression; the hybrid replicated well in both COP-5 and

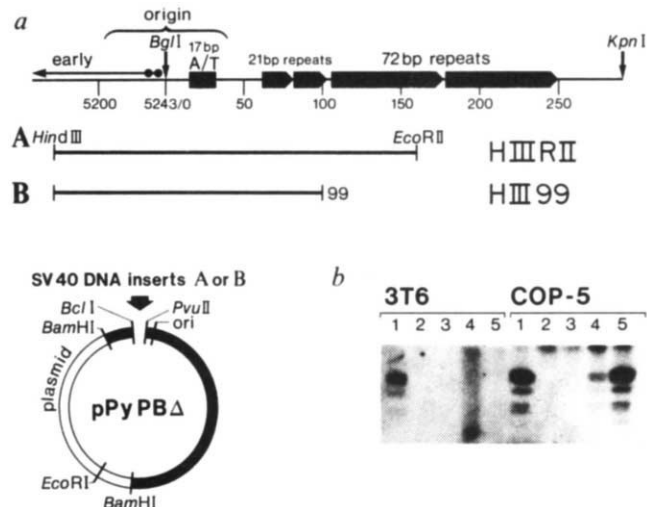


Fig. 2 DNA replication assays in 3T6 and COP-5 cells. **a**, Schematic representation of SV40 DNA fragments inserted into pPyPBΔ. These fragments were obtained from recombinants provided by Yakov Gluzman of Cold Spring Harbor. **b**, Autoradiograph showing replication of the DNAs used to transfect the cells: lane 1, pPy (wild type A2 plasmid; Table 1); lane 2, pPyPBHII99- (pPyPBΔ with insert B, in the opposite orientation to that shown); lane 3, pPyPBHIIIRII- (pPyPBΔ with insert A, in the opposite orientation to that shown); lane 4, pPyPBHII99+ (insert B, in orientation shown); lane 5, pPyPBHIIIRII+ (insert A, in orientation shown).

Methods: The recombinant DNAs were not excised from the plasmid vector before transfection and the cells were transfected by the calcium phosphate coprecipitation technique, as described in detail elsewhere⁴⁸. Otherwise the replication assays were done as described in Fig. 1 legend.

3T6 cells (Fig. 1, lanes 4; Table 1). The 21-bp repeats are important for both SV40 early transcription and DNA replication^{1,33-35,55,56}, whereas the 72-bp repeat sequence (plus some adjacent sequences on the late side) comprises the SV40 transcriptional enhancer^{4,6}. The results obtained when we assayed replication of the hybrid DNAs (Fig. 1) demonstrated consistently that progressive deletion of the SV40 sequences from the *HindIII* site, removing first the SV40 ORI and then the 21-bp repeats, did not significantly decrease replication of the hybrid DNAs in COP-5 cells. A single copy of the 72-bp repeat (Fig. 1, fragment E) was sufficient to activate polyoma virus DNA replication in both 3T6 and COP-5 cells (Table 1). All these fragments functioned efficiently when inserted in either orientation, but fluctuations among experiments discouraged precise quantitative comparisons between constructs. A hybrid containing only 54 nucleotides of a single 72-bp repeat (Fig. 1, fragment F) could not replicate in 3T6 cells (Fig. 1, lanes 8; Table 1) but was functional in COP-5 cells, in which replicatory but not transcriptional function was assayed. Interestingly, a dimer insert of the same fragment did allow replication in 3T6 cells (Fig. 1, lanes 9). We directly tested the implication that a dimer insert of a 'crippled' SV40 enhancer could restore its function, by staining for polyoma virus T-antigen immunofluorescence in 3T6 and HeLa cells transfected with the appropriate hybrids. T-antigen expression by the hybrid with a monomer insert was extremely inefficient, while a dimer insert increased this level of expression by at least 10-fold (Table 1).

To confirm that the 72-bp repeat sequences represent the element within the SV40 regulatory region which restores both transcriptional and replicatory ability to pPyPBΔ, we also analysed fragments deleted from the SV40 *KpnI* site (Fig. 2). The SV40 fragment containing the ORI plus 21-bp repeats (Fig. 2, fragment B) was extremely inefficient in restoring polyoma virus DNA replication in COP-5 cells when compared with wild-type DNA and even with fragments containing only partial sequences from the 72-bp repeat (Fig. 1, lane 8; Fig. 2, lanes 4, 5). This

Table 1 Summary of replication and expression results

Recombinant	DNA fragment inserted into pPyPBA	DNA replication in		T antigen immunofluorescence in HeLa cells
		3T6	COP-5	
pPyPBHIIK+ (1× or 2×)	Fig. 1 fragment A	++	++	>20% +ve
pPyPB10K (+ or -, 1× or 2×)	Fig. 1 fragment B	++	++	ND
pPyPB48K (+ or -, 1× or 2×)	Fig. 1 fragment C	++	++	ND
pPyPB98K (+ or -, 1× or 2×)	Fig. 1 fragment D	++	++	>20% +ve
pPyPB164K (+ or -, 1× or 2×)	Fig. 1 fragment E	++	++	>20% +ve
pPyPB196K+	Fig. 1 fragment F	-	+	≤0.1% +ve
pPyPB196K(2×)+	Dimer insert F	+	++	~2% +ve
pPyPBHIIIRII+	Fig. 2 fragment A	-	+	*
pPyPBHIIIRII-	Fig. 2 fragment A (opposite orientation)	-	-	ND
pPyPBHIIIR99+	Fig. 2 fragment B	-	±	†
pPyPBHIIIR99-	Fig. 2 fragment B (opposite orientation)	-	-	ND
pPyPBPyHIIIRI	Polyoma virus coding DNA (1,560-1,656)	-	-	ND
pPyPBSVHIIIRI	SV40 coding DNA (1,708-1,782)	-	-	ND
pPyPBADMLpro	Adenovirus-2 major late promoter (-260 to +93)	-	-	ND
pPyPBβpro	Rabbit β-globin-1 gene promoter (-109 to -10)	-	-	ND
'Parental' constructs				
pPy	Strain A2 polyoma virus DNA inserted into <i>Bam</i> HI site of pAT153	++	++	4-6% +ve
pPyPBA	pPy with the <i>Bcl</i> II to <i>Pvu</i> II fragment from nucleotide 5,021 to 5,265 replaced by a <i>Xho</i> I linker	-	-	<0.001%

Polyoma virus and SV40 numbering according to Tooze³¹. pAT153 is a derivative of pBR322 (ref. 49). + and -, Orientations of the inserts (Figs 1, 2). 1× and 2×, Monomer and tandem dimer inserts, respectively. All of the replication assays were repeated at least three times, using different transfection methods and assaying replication at different times after transfection; the results were always consistent. ++, Replication approximately equivalent to that of polyoma virus DNA in all assays. +, Constructs which, although replicating to wild type polyoma virus DNA levels in a 2.5-day period, replicate significantly less efficiently than wild type if assayed after only 24 h (for example, replication in COP-5 cells, Fig. 1, lane 8 compared with Fig. 4, lane 5). ±, Only weak replication detectable after 2.5 days (Fig. 2, lane 4). -, No replication detectable even after 2.5 days. After 2.5 days, transfected cells were fixed and stained as described previously⁵⁰, 8×8 mm areas, containing 4-5×10⁴ cells, were stained and the total numbers of cells and proportion positive for nuclear T antigen were extrapolated after precisely counting areas of 0.145 mm². Similar results were obtained using 3T6 cells (data not shown). ND, not done.

* Not tested here, but the same fragment within SV40 gives 2-5 % of wild-type numbers of T antigen-positive HeLa cells.

† Yields up to 1 % of wild-type SV40 T antigen levels in HeLa cells.

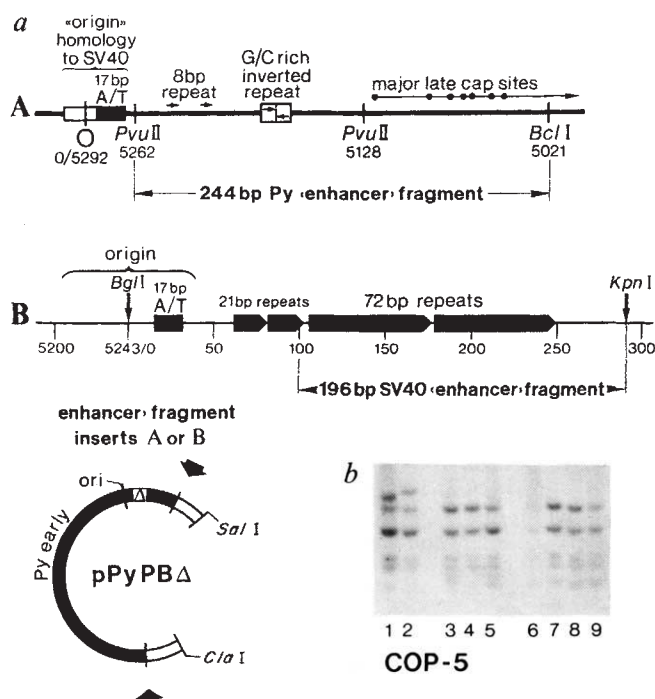


Fig. 3 DNA replication assay in COP-5 cells. *a*, Schematic representation of the hybrid recombinant DNAs used. The 244-bp polyoma virus enhancer fragment (A) and the 196-bp SV40 enhancer fragment (B) were subcloned, via *Xho*I linkers, into the 3' noncoding region of the rabbit β-globin gene²³. We then excised ~1-kb *Bgl*II restriction fragments from each, which have the respective viral enhancers flanked on either side by ~400 bp of rabbit DNA sequences, and inserted these into the two *Bam*HI sites of the polyoma virus recombinant pPyPBA (arrows). *b*, Autoradiograph showing replication in COP-5 cells: lane 1, pPy (Fig. 2, Table 1); lane 2, pPyPBHIIK+ (Fig. 1); lane 3, pPyPBA with insert B at the proximal site; lane 4, pPyPBA with insert B at the distal site; lane 5, pPyPBA with insert B at the distal site (in the opposite orientation to that shown); lane 6, insert A at the proximal site (some material was lost in this lane and other experiments revealed that this construct replicates as efficiently as the others); lane 7, pPyPBA with insert A at the proximal site (opposite orientation); lane 8, pPyPBA with insert A at distal site (+ orientation); lane 9, pPyPBA with insert A at distal site (- orientation). **Methods:** Hybrid viral DNAs were excised from the bulk of plasmid sequences by *Clal*/*Sal*I double digestions. Linear molecules were used to transfect 3T6 cells (not shown) and COP-5 cells, using DEAE-dextran, as before. Hirt supernatants were prepared after 2.5 days and DNA replication assayed as described in Fig. 1 legend.

was surprising as we expected to find a role for the 21-bp repeats in DNA replication. The SV40 fragment comprising the ORI, the 21-bp repeats and only 53 nucleotides of the origin-proximal 72-bp repeat (Fig. 2, fragment A) was able to restore replication in only one orientation (Fig. 2, lanes 3, 5). In the other orientation the SV40 early promoter lies between the SV40 enhancer sequences and the polyoma virus origin of replication, presumably blocking the enhancer function. The effects of intervening promoters on enhancer function has been well documented^{21,22,25} and this supported our hypothesis that it is enhancer function that is required in *cis* to activate polyoma virus DNA replication.

Constructs containing SV40 enhancer fragments (Fig. 1, inserts A, D, E) were further assayed for transcriptional enhancer activity by S₁ nuclease mapping of 5' termini after transient expression in 3T6 and HeLa cells. In all cases we found correctly initiated polyoma virus early mRNA independently of the insert orientation (unpublished data). T-antigen immunofluorescence staining experiments confirmed the indirect expression data obtained by measuring DNA replication in 3T6 cells (Table 1). We also found that sequences which were less efficient in restoring DNA replication in COP-5 cells had corresponding partial abilities to restore viral early gene expression. We did not observe DNA replication in either 3T6 or COP-5 cells after insertion of other heterologous sequences derived from protein coding regions or from known transcriptional promoters (Table 1).

A further parallel between the functional properties of transcriptional enhancers and those of sequences activating DNA replication was observed when we assayed replication of additional recombinant hybrids (Fig. 3). These constructs have either the SV40 enhancer or the polyoma virus enhancer inserted into the deleted plasmid pPyPBA at two distant sites (~800 and 1,400 bp from the ORI, respectively). After deletion of plasmid DNA sequences, these inserts efficiently restored DNA replication (Fig. 3) when placed in either orientation at either position. The distant inserts also restored early gene expression, as measured by detection of DNA replication in 3T6 cells and by immunofluorescence staining for T antigen in HeLa cells (data not shown).

Cell-type specific activation of replication

Viral enhancers function more efficiently in certain cell types than in others^{8,10,23,35,36}. We have reported replication of SV40 enhancer-polyoma virus hybrid DNAs in human HeLa cells, after replacing the polyoma virus enhancer with the enhancer region of SV40^{21,23}. Recently, DNA sequences with transcriptional enhancer function have been isolated from cellular DNA and some of these exhibit extreme tissue specificity. The first such enhancer was identified within an intervening sequence of the mouse immunoglobulin heavy-chain gene. The immunoglobulin enhancer functions efficiently in cells of myeloid origin but hardly, if at all, in fibroblasts¹⁵⁻¹⁷.

We tested this known cellular transcriptional control element for its ability to activate the replication of an enhancer-deleted polyoma virus plasmid. There is no *a priori* reason to expect the immunoglobulin enhancer coincidentally to be an activator of viral DNA replication, unless the two functions are inherent properties of a single class of control elements. We replaced the polyoma virus regulatory region with a fragment of cellular DNA known to contain the tissue-specific immunoglobulin enhancer (Fig. 4). The resulting hybrid was assayed for DNA replication, both in the fibroblast-derived COP-5 cell line and in the J558L³⁷ myeloma line. No replication of the immunoglobulin enhancer-polyoma virus hybrid was detected in COP-5 cells, even though these cells provide T antigen. Control molecules containing either the wild-type polyoma virus sequence or the SV40 enhancer, or two copies of a truncated SV40 enhancer, all replicated well. In contrast, the immunoglobulin enhancer-polyoma virus hybrid replicated more efficiently than the control molecules in the J558L cells (Fig. 4: in this experiment we detected replicated DNA by its

acquired resistance to digestion by restriction endonuclease *DpnI* which only cleaves methylated DNA³⁸). The observation that the immunoglobulin recombinants replicates more efficiently than the control constructs is somewhat surprising because both the polyoma virus and SV40 enhancers function in myeloma cells¹⁵. Further experiments are in progress to determine whether this effect is *cis*-specific for DNA replication or a consequence of differential large-T antigen expression. We obtained the same results with a second myeloma cell line, X63-Ag8 (unpublished data). These experiments demonstrate that polyoma virus DNA replication can be activated by sequences otherwise identified only as transcriptional enhancers.

Conclusions

The polyoma virus transcriptional enhancer is of considerable interest because there are sequences within this region of the viral genome that are also required to activate DNA replication, and we have sought to resolve and separate these functions. We found that only enhancer sequences were able to substitute effectively for either the transcriptional or replicatory activator function. The only separation we obtained appears quantitative. Although truncated enhancer fragments from SV40 were able to activate DNA replication effectively, they were extremely inefficient in the restoration of early gene expression. Significantly, dimerization of such a truncated fragment dramatically increased its ability to enhance gene expression. The observations that viral enhancer sequences could activate DNA replication in an orientation- and position-independent manner led us to suggest that the polyoma virus transcriptional enhancer

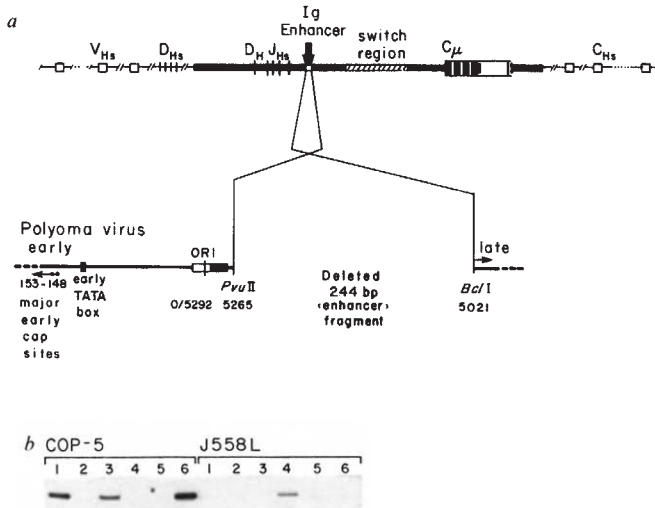


Fig. 4 Replication assays in COP-5 and myeloma J558L cells. *a*, Schematic representation of the mouse heavy-chain immunoglobulin gene (adapted from Banerji *et al.*¹⁵), showing the 303-bp *PstI* to *EcoRI* DNA fragment, bearing the immunoglobulin enhancer function¹⁵, which was inserted via synthetic DNA linkers into pPyPBA. *b*, Autoradiographs showing replication of input plasmid DNAs in COP-5 and J558L cells: lane 1, pPy (Fig. 2, Table 1); lane 2, pPyPBA (Fig. 1); lane 3, pPyPB164K- (see Fig. 1, pPyPBA with insert E); lane 4, pPyPB164K+ (pPyPBA with immunoglobulin enhancer insert); lane 5, pPyPB196K+ (Fig. 1, insert F); lane 6, pPyPB196K(2x)+ (Fig. 1, dimer insert F).

Methods: Unexcised plasmid DNAs were used to transfect cells with DEAE-dextran as before. Hirt supernatants were prepared from lysed cells after only 24 h and, in order to increase the sensitivity of the assay, we digested the lysates with *XbaI* then *DpnI* (which cleaves the input DNA at multiple sites but cannot cleave the newly replicated DNA)³⁸, generating ~9-kb linear molecules from replicated DNA. We blotted these to nitrocellulose filters after gel electrophoresis, as before, and used T4 DNA polymerase to label *Sau3A* fragments of pPyPBA with ³²P for use as a hybridization probe.

and the replicatory activator are one and the same. Consistent with this hypothesis was the observation that activation of DNA replication appeared to be inhibited by the intervention of the SV40 early promoter between the enhancer sequences and the polyoma virus replication origin.

The identification of the immunoglobulin heavy-chain gene enhancer¹⁵⁻¹⁷ allowed us to test the hypothesis definitively and we found that polyoma virus DNA replication is activated by enhancers of both viral and cellular origin. The observation that the immunoglobulin enhancer-polyoma virus hybrid DNA replicates efficiently in myeloma cells lends support to our previous proposal that viral enhancers have an important role in determining the host ranges of viruses²³.

It is not entirely unexpected that enhancer function is involved directly in polyoma virus DNA replication. Isolated variants of polyoma virus which are able to grow in embryonal carcinoma cells³⁹⁻⁴¹ all have sequence alterations within the enhancer region of the viral genome. Although the sequence changes appear to increase the efficiency of the viral enhancer^{25,42}, they act only in *cis*; such mutants cannot complement the growth of wild-type virus in embryonal carcinoma cells²⁶.

Comprehensive deletion analysis of the polyoma virus enhancer region has identified a 23-bp segment with minimal replicatory activator ability. Tandem duplication of this fragment restores full activation of DNA replication while further oligomerization produces an efficient transcriptional enhancer (T. Veldman, S.L. and R.K., in preparation).

There are several possible explanations for the role of enhancers in DNA replication. For instance, if enhancers function by localizing DNA at the nuclear matrix or by altering the conformation of the template^{21,24}, one could envisage that such molecular mechanisms could equally be involved in DNA replication or transcription. A further possibility is that transcription is directly involved in triggering DNA replication. Although there is no direct experimental support for this proposal, the initiation of polyoma virus DNA replication involves RNA primers⁴³ and clear precedents for transcriptional activation of DNA synthesis exist in prokaryotic systems^{44,45}. There may be

a switch of early gene transcription initiation sites at about the time DNA replication begins, in both polyoma virus and SV40^{46,47}, and such a T antigen-mediated switch mechanism could be involved in triggering DNA replication²⁵.

Clearly, enhancers are not always essential for viral DNA replication. The most readily comparable and extensively studied example is SV40 DNA replication. Enhancer function is not required for the replication of SV40 DNA, and although the 21-bp repeats have a clear but non-essential role in replication^{33,34}, no involvement of the 72-bp repeat enhancer region has yet been shown. This discrepancy could be explained by the differences around the origins of DNA replication of these two related viruses. The SV40 early promoter is itself integrated into the SV40 ORI sequences. The A+T-rich nucleotide stretch, which forms part of the ORI sequences conserved between papova viruses, apparently serves as the 'TATA box' for the SV40 major early promoter⁴⁸. If transcriptional activation is involved in DNA replication, the presence of this strong promoter within the ORI could circumvent an enhancer dependence. It is also possible that the equivalent A+T-rich stretch in the polyoma virus ORI is part of a replication-activating promoter.

Our experiments indicate that T antigen-dependent replication is possible in specialized cell types, an enhancer is essential in *cis* (in the presence of provided T antigen) to activate the polyoma virus ORI and also that a cellular enhancer, or even a 'crippled' enhancer, can suffice. This system could therefore be adapted to selectively isolate other cellular enhancers (including tissue-specific enhancers) from randomly generated DNA fragments. It remains to be seen whether enhancers play a general role in the functioning of replication origins, but we speculate that cellular DNA replication could be coordinated by inducible enhancers.

We thank Julian Banerji and Trudi Veldman for valuable discussions and Marybeth Erker, Martha Richardson, Silvia Oberholzer and Fritz Ochsenbein for their help in the preparation of this text. The work was supported by the Swiss NRF, the Kanton of Zurich, the ICRF and Genetics Institute, Inc.

Received 21 June; accepted 28 August 1984.

- Benoist, C. & Chambon, P. *Nature* **290**, 304-310 (1981).
- Gruss, P., Dhar, R. & Khoury, G. *Proc. natn. Acad. Sci. U.S.A.* **78**, 943-947 (1981).
- Tyndall, C. *et al. Nucleic Acids Res.* **9**, 6231-6250 (1981).
- Banerji, J., Rusconi, S. & Schaffner, W. *Cell* **27**, 299-308 (1981).
- de Villiers, J. & Schaffner, W. *Nucleic Acids Res.* **9**, 6251-6264 (1981).
- Moreau, P. *et al. Nucleic Acids Res.* **9**, 6047-6068 (1981).
- Lusky, M. *et al. Molec. cell. Biol.* **3**, 1108-1122 (1983).
- Laimins, L. A. *et al. Proc. natn. Acad. Sci. U.S.A.* **79**, 6453-6457 (1982).
- Jolly, D. J. *et al. Nucleic Acids Res.* **11**, 1855-1872 (1983).
- Kriegler, M. & Botchan, M. *Molec. cell. Biol.* **3**, 325-339 (1983).
- Luciw, P. A. *et al. Cell* **33**, 705-716 (1983).
- Hearing, P. & Shenk, T. *Cell* **33**, 695-703 (1983).
- Hen, R. *et al. Nucleic Acids Res.* **11**, 8747-8760 (1983).
- Cordingley, M. G., Campbell, M. E. M. & Preston, C. M. *Nucleic Acids Res.* **11**, 2347-2365 (1983).
- Banerji, J., Olson, L. & Schaffner, W. *Cell* **33**, 729, 740 (1983).
- Gillies, S. D. *et al. Cell* **33**, 717-728 (1983).
- Neuberger, M. S. *EMBO J.* **2**, 1373-1378 (1983).
- Picard, D. & Schaffner, W. *Nature* **307**, 80-82 (1984).
- Queen, C. & Baltimore, D. *Cell* **33**, 741-748 (1983).
- Walker, M. D. *et al. Nature* **306**, 557-561 (1983).
- de Villiers, J. *et al. Cold Spring Harb. Symp. quant. Biol.* **47**, 911-919 (1983).
- Wasyluk, B. *et al. Cell* **32**, 503-514 (1983).
- de Villiers, J. *et al. Nucleic Acids Res.* **10**, 7965-7976 (1982).
- Khouri, G. & Gruss, P. *Cell* **33**, 313-314 (1983).
- Gluzman, Y. & Shenk, T. *Enhancers and Eukaryotic Gene Expression* (Cold Spring Harbor Laboratory, New York, 1983).
- Fujimura, F. K. & Linney, E. *Proc. natn. Acad. Sci. U.S.A.* **79**, 1479-1483 (1982).
- Luthman, H., Nilsson, M.-G. & Magnusson, G. *J. molec. Biol.* **161**, 533-550 (1982).
- Muller, W. J. *et al. J. Virol.* **47**, 586-599 (1983).
- Katinka, M. & Yaniv, M. *J. Virol.* **47**, 244-248 (1983).
- Treizenberg, S. J. & Folk, W. *J. Virol.* **51**, 437-444 (1984).
- Toozie, J. *DNA Tumor Viruses 2nd edn* (Cold Spring Harbor Laboratory, New York, 1981).
- Geier, G. E. & Modrich, P. *J. biol. Chem.* **254**, 1408-1413 (1979).
- Fromm, M. & Berg, P. *J. molec. appl. Genet.* **1**, 457-481 (1982).
- Bergsma, D. J. *et al. Proc. natn. Acad. Sci. U.S.A.* **79**, 381-385 (1982).
- Byrne, B. J. *et al. Proc. natn. Acad. Sci. U.S.A.* **80**, 721-725 (1983).
- Spandidos, D. A. & Wilkie, N. M. *EMBO J.* **2**, 1193-1199 (1983).
- Oi, V. *et al. Proc. natn. Acad. Sci. U.S.A.* **80**, 825-829 (1983).
- Lacks, S. & Greenberg, B. *J. biol. Chem.* **250**, 4060-4066 (1975).
- Katinka, M. *et al. Cell* **20**, 393-399 (1980).
- Fujimura, F. K. *et al. Cell* **23**, 809-814 (1981).
- Sekikawa, K. & Levine, A. J. *Proc. natn. Acad. Sci. U.S.A.* **78**, 1100-1104 (1981).
- Linney, E. *et al. in Teratocarcinoma Stem Cells* (eds Silver, L. M., Martin, G. R. & Strickland, S.) 271-284 (Cold Spring Harbor Laboratory, New York, 1983).
- Magnusson, G. *et al. Proc. natn. Acad. Sci. U.S.A.* **70**, 412-420 (1973).
- Dove, W. F., Inokuchi, H. & Stevens, W. F. *in The Bacteriophage Lambda* (ed. Hershey, A. D.) 747-777 (Cold Spring Harbor Laboratory, New York, 1971).
- Panayotatos, N. *Nucleic Acids Res.* **12**, 2641-2648 (1984).
- Fenton, R. G. & Basilico, C. *Proc. natn. Acad. Sci. U.S.A.* **79**, 7142-7146 (1982).
- Ghosh, P. K. & Lebowitz, P. *J. Virol.* **40**, 224-240 (1981).
- Ghosh, P. K. *et al. Proc. natn. Acad. Sci. U.S.A.* **78**, 100-104 (1981).
- Twigg, A. J. & Sherratt, D. *Nature* **283**, 216-218 (1980).
- de Villiers, J. & Schaffner, W. *in Techniques in Nucleic Acid Biochemistry* (ed. Flavell, R. A.) B507, 1-20 (Elsevier, Amsterdam, 1984).
- Legerski, R. J., Hodnett, J. L. & Gray, H. B. *Nucleic Acids Res.* **5**, 1445-1464 (1978).
- Maniatis, T., Fritsch, E. F. & Sambrook, J. *Molecular Cloning (A Laboratory Manual)* (Cold Spring Harbor Laboratory, New York, 1982).
- O'Hare, K. *J. molec. Biol.* **151**, 203-210 (1981).
- Hirt, B. *J. molec. Biol.* **26**, 365-369 (1967).
- Everett, R. D., Baty, D. & Chambon, P. *Nucleic Acids Res.* **11**, 2447-2464 (1983).
- Baty, D. *et al. Nucleic Acids Res.* **12**, 915-932 (1984).

Role of DNA replication in the repression of silent mating type loci in yeast

Allan M. Miller & Kim A. Nasmyth

MRC Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, UK

A putative origin of DNA replication is associated with the DNA sequences necessary for the repression of silent mating type loci in yeast. These sequences lie about a kilobase away from the affected promoters, so the repression must act at a distance. We show here that DNA replication is required for the onset of repression.

IN the yeast *Saccharomyces cerevisiae*, mating type is determined by the *MAT* locus, which has two alleles, *a* and α . *MATa* and *MAT α* differ in a 700-base pair (bp) region which includes the promoter and part of the coding sequences of two divergent *a*- or α -specific transcripts^{1,2}. Yeast strains that carry the wild type *HO* gene are able to switch mating types by the directed mutation of one *MAT* allele to the other³, which is achieved by replacing the *a*- (or α -) specific sequence at *MAT* with the α - (or *a*-) specific sequence. Physical and genetic analysis of this process has demonstrated that there exist unexpressed copies of mating type DNA which act as donors of sequence information in the switching process (for review, see ref. 4). These silent copies, denoted *HML α* and *HMRa*, contain the complete transcription units found at *MAT α* and *MATa* respectively, but are not transcribed due to their different flanking sequences^{2,5}.

Deletion analysis of the silent copies has demonstrated that *HML* and *HMR* are under the negative control of their flanking DNA sequences⁶. A small segment of DNA about a kilobase (kb) to the left of each silent copy promoter (called *E_L* for *HML* and *E_R* for *HMR*) is essential for their repression. Also required for the repression of both silent copies are four *trans*-acting gene products *SIR 1, 2, 3* and *4* (refs 7–10).

Several observations suggest that one can consider *E* as a negative enhancer¹¹. First, *E* can repress a promoter up to 2.5 kb

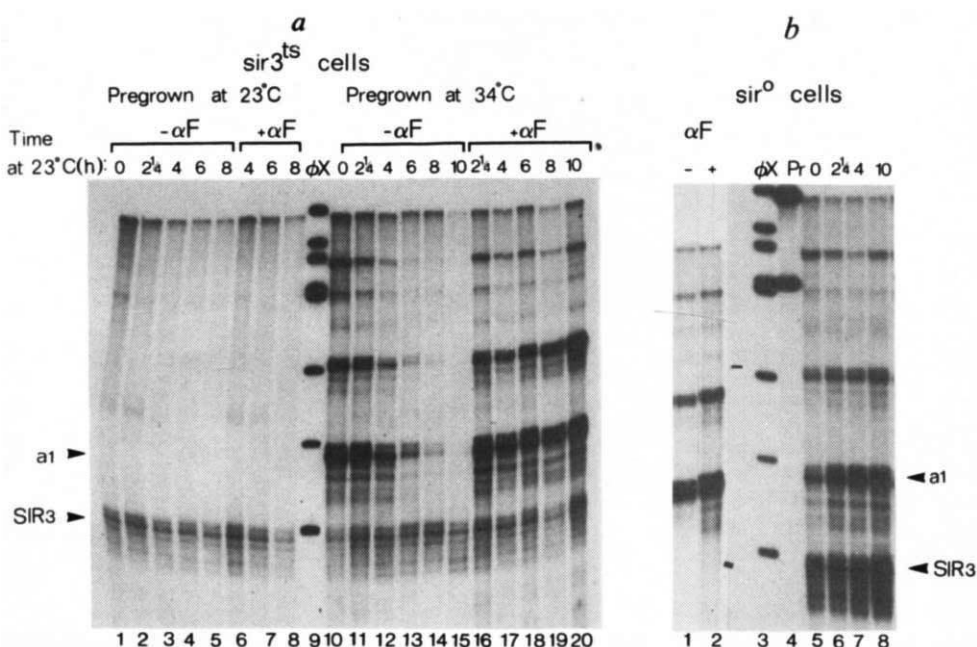
away. Second, *E* can repress other promoters when substituted for the *a* promoter at *HMR*. Third, the relative orientations of both *E* and the regulated promoter are unimportant (A. Brand and L. Breeden, personal communication).

The observation¹² that the *E* sequence is very tightly linked to an autonomously replicating sequence (ARS) suggests that DNA replication may be important in *SIR*-mediated repression (ARSs confer the ability to replicate autonomously and are probably specific origins of DNA replication^{13–17}). Here we test whether DNA replication is involved in *SIR* repression. More specifically, we investigate the expression of *HML* and *HMR* in strains carrying *ts* (temperature-sensitive) alleles of *SIR3* or *SIR4* and ask whether DNA replication is required for repression or for derepression of *HML* and *HMR* in these strains. Our results suggest that DNA replication is required for the onset of repression (after a shift to the permissive temperature), whereas depression (after a shift to the restrictive temperature) can occur in the absence of cell cycle progress.

Repression of *HML* and *HMR*

To test whether DNA replication is required for *SIR* repression, we tested whether the onset of *SIR* repression can occur in growing cells in the absence of any cell cycle progress. We used strain M28 (*HMLa MAT Δ HMRa sir3^{ts}*) (Table 1), which

Fig. 1 *a*, Levels of *a1* RNA in strain M28 after shift to the permissive temperature. Strain M28 (Table 1) was grown into G₀ and purified daughter cells were inoculated into fresh medium at 23 °C in the presence or absence of 10 units ml⁻¹ α -factor. Cells were collected at intervals and the level of *a1* RNA was determined by the protection of radiolabelled DNA from S₁ digestion. DNA probe homologous to the 3' end of the *SIR3* transcript was mixed with the *a1* probe to act as an internal control. 25 μ g of total RNA were used per reaction. Lanes 1–8, cells grown into G₀ at the permissive temperature (23 °C). Lane 9, Φ X174 $\times$ *Hinf*I. Lanes 10–20, cells grown into G₀ at the restrictive temperature (34 °C). The time elapsed (h) since inoculation is shown. This experiment yields the same result if purified mothers or total G₀ cells are used. Daughters are used here to allow more *SIR3* synthesis at the permissive temperature in the α -factor blocked cells. *b*, Transcription from the *a1* promoter is not affected by α -factor or by cell cycle progress. Strain M29 (Table 1) is isogenic with M28, except that it is *sir3⁰*, and therefore expresses *HML* and *HMR* constitutively. The effect of α -factor on the level of *a1* RNA: Cells of strain M29 were grown into G₀ at 34 °C and inoculated into fresh YEPD medium at 23 °C. After 7 h growth, RNA samples were taken. Lane 1, cells grown without α -factor. Lane 2, cells grown in the presence of 10 units ml⁻¹ α -factor. Lane 3, Φ X174 $\times$ *Hinf*I. Lane 4, undigested probe (Pr). 40 μ g of total RNA were used per reaction. The effect of cell cycle progress on the level of *a1* RNA: Cells of strain M29 were grown to G₀ at 34 °C and purified daughters were inoculated into fresh YEPD at 23 °C. RNA samples were taken at intervals: lane 5, 0 h; lane 6, 2½ h; 7, 4 h; 8, 10 h. 25 μ g total RNA were used per reaction. Daughter cell purification, RNA preparation and S₁ protection were performed as described previously³⁶. α -factor was prepared according to the procedure of Bucking-Throm *et al.*³⁷.



transcription from the *a1* promoter is not affected by α -factor or by cell cycle progress. Strain M29 (Table 1) is isogenic with M28, except that it is *sir3⁰*, and therefore expresses *HML* and *HMR* constitutively. The effect of α -factor on the level of *a1* RNA: Cells of strain M29 were grown into G₀ at 34 °C and inoculated into fresh YEPD medium at 23 °C. After 7 h growth, RNA samples were taken. Lane 1, cells grown without α -factor. Lane 2, cells grown in the presence of 10 units ml⁻¹ α -factor. Lane 3, Φ X174 $\times$ *Hinf*I. Lane 4, undigested probe (Pr). 40 μ g of total RNA were used per reaction. The effect of cell cycle progress on the level of *a1* RNA: Cells of strain M29 were grown to G₀ at 34 °C and purified daughters were inoculated into fresh YEPD at 23 °C. RNA samples were taken at intervals: lane 5, 0 h; lane 6, 2½ h; 7, 4 h; 8, 10 h. 25 μ g total RNA were used per reaction. Daughter cell purification, RNA preparation and S₁ protection were performed as described previously³⁶. α -factor was prepared according to the procedure of Bucking-Throm *et al.*³⁷.

Table 1 Genotypes of the strains used

Strain	Genotype
M9	<i>HMLα MATα HMRa sir3^{ts} cdc28-4^{ts} ade2 leu2 met2 tyr1 ura1</i>
M10	<i>HMLα MATα HMRa sir3^{ts} CDC⁺ ade2 leu2 met2 ura1</i>
M20	<i>HMLα MATα HMRa sir3^{ts} cdc28-4^{ts} ade2 met2 trp1 tyr1</i>
M26	<i>HMLα MATα HMRa sir3^{ts} cdc28-4^{ts} ade2 met2 TRP1::SIR3 tyr1</i>
M28	<i>HMLα MATΔ HMRa sir3^{ts} trp1 ura1 ade2 his4</i>
M29	<i>HMLα MATΔ HMRa sir3⁰::TRP1 trp1 ura1 ade2 his4</i>
M37	<i>HMLα MATΔ HMRa sir4^{ts} ade2 can1-100</i>

Strain M26 is isogenic with strain M20 but contains the *SIR3* gene integrated at the *TRP1* locus, and strain M29 is isogenic with M28. The *sir3^{ts}* and *sir4^{ts}* alleles used are *ste8-59a* and *ste9-62c* respectively³⁴. *cdc28-4* is from S. I. Reed³⁵. *sir3⁰*::*TRP1* denotes the insertion of the *TRP1* gene into the open reading frame of the *SIR3* gene.

should produce *a1* transcripts from both of its silent copies at the restrictive temperature, but none at all at the permissive temperature. Because the *MATa1* primary transcript is processed and the mature mRNA produced has a short half life ($t_{1/2} < 3$ min)¹⁸, we can use the level of the mature RNA or of its precursors as a measure of the instantaneous rate of transcription. Cells of M28 grown into stationary phase (G_0) at the restrictive temperature transcribe *HMLa1* and *HMRa1* fully. We then analysed the rate of *a1* mRNA transcription after these cells were inoculated into fresh medium at the permissive temperature and tested whether it is affected by blocking cell cycle progress with the yeast pheromone α -factor (α -factor blocks *a* cells at early G_1 but does not directly affect macromolecular synthesis)¹⁹.

In the absence of α -factor, we find that repression (measured by a reduction in the level of *a1*) starts between 2.5 and 4 h and is virtually complete by 10 h. In the presence of α -factor, however, no repression occurs even after 10 h (Fig. 1a). The effect of α -factor cannot be due to some gross inhibition of cell growth, since a significant degree of repression occurs within 4 h (at which time α -factor has little or no effect on overall cell growth (A_{660}) or RNA synthesis) (Fig. 2).

This inhibition of *a1* mRNA disappearance suggests that cell cycle progress is required for the repression of previously derepressed copies of *HML* and *HMR*. This result could, alternatively arise if *SIR3* synthesis or the activity of the *a1* promoter were directly affected by α -factor or by cell cycle progress. Figure 1a shows that the level of *SIR3* transcript is unaltered throughout the experiment, both in the presence and absence of α -factor. The effect of these treatments on the activity of the *a1* promoter can be followed in strain M29, which is isogenic with M28 but is *sir3⁰* (Table 1). (*sir3⁰* denotes a *SIR3* gene which has been inactivated by the insertion of another gene into its open reading frame.) When G_0 cells from M29 are inoculated into fresh medium (with or without α -factor), no significant change in the level of *a1* is seen (Fig. 1b), demonstrating that the *a1* promoters at the silent copies are not affected by these treatments.

The changes observed in the level of *a1* transcript cannot be due to a change in its half life, as we can follow the processing of this message. The *MATa1* gene is spliced and the larger fragments observed correspond to unspliced precursor and what is probably a splicing intermediate¹⁸; the coordinate drop in the steady-state levels of all these species must therefore reflect a change in the rate of transcription.

Finally, α -factor does not in itself cause derepression of the silent copies. M28 cells grown to G_0 at the permissive temperature and inoculated into fresh medium (with or without α -factor) at the permissive temperature show no silent copy transcription (Fig. 1a). When all the experiments shown in Fig. 1 are repeated using exponentially growing cells, the same results are observed. The effect of α -factor is therefore not peculiar to cells grown out of G_0 .

Thus, α -factor does not directly affect *a1* or *SIR3* transcription, and neither of these transcripts shows any cell cycle depen-

dence, yet α -factor blocks the onset of repression of both *HMLa* and *HMRa*. As this effect is observable before any difference in overall macromolecular synthesis, we conclude that progress through the cell cycle is necessary for *SIR*-mediated repression.

Timing of repression

We have shown that *SIR* repression cannot occur in cells blocked in G_1 by α -factor, so it must occur at some later stage in the cell cycle. Figure 3a shows that when cells are released from the α -factor block, repression subsequently occurs. The kinetics of repression during this release are similar to those of bud emergence, suggesting that it takes place around S phase.

To determine more precisely the cell cycle events on which repression is dependent, we tested whether the repression which occurs upon release is blocked either by an inhibitor of DNA synthesis (hydroxyurea²⁰) or by an inhibitor of mitosis (methylbenzimidazol-2-yl carbamate^{21,22}, MBC). We find that hydroxyurea blocks repression while MBC does not (Fig. 3b). Repression therefore takes place at a stage after the hydroxyurea block but before the MBC block, suggesting that DNA replication itself is required. The same result is obtained when this experiment is repeated using a *sir4^{ts}* allele (strain M37 of Table 1) (data not shown), implying that the dependence on cell cycle progress for repression is not a peculiarity of the *sir3^{ts}* allele, but rather reflects a property of both the *SIR3* and *SIR4* gene products.

DNA replication not required for derepression

We have determined that if *HML* and *HMR* become derepressed, then the *SIR⁺* gene products can only re-exert repression during S phase. We now ask whether *SIR⁺* activity is required throughout the cell cycle to maintain repression, or whether *SIR* acts solely at DNA replication. *sir3^{ts}* cells were grown to G_0 at the permissive temperature and inoculated into fresh medium at the restrictive temperature. In this case, cell cycle progress can be blocked either by α -factor or by a *ts* allele of *CDC28* (*CDC28* is required for progress through G_1 ; ref. 23).

Transcription from *HMRa* is induced within 1 h of inoculation at the restrictive temperature, irrespective of whether the cells were blocked in G_1 (M20 *sir3^{ts} cdc28^{ts}*) or not (M10 *sir3^{ts} CDC⁺*) (Fig. 4a). Identical results are obtained using M28 in

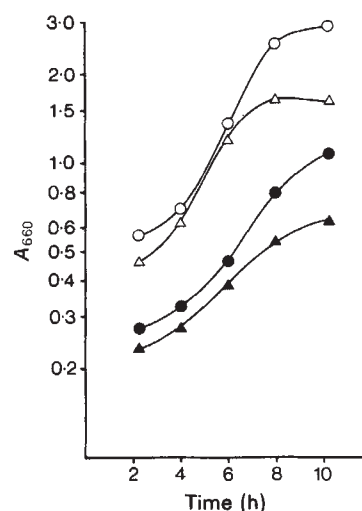


Fig. 2 Growth curves. Graph of the absorbance of the culture and of the total RNA yield. The A_{660} of the culture used in Fig. 1a is plotted against time for the two experimental conditions: ▲, α -factor; and ●, no α -factor. After a short lag phase, the absorbance increases exponentially. The decrease in rate at 8 h is an artefact caused by the nonlinearity of the absorbance readings. Also shown is the yield of total RNA obtained in the RNA preparations: △, plus α -factor; and ○, minus α -factor (for the cells that were grown to G_0 at the restrictive temperature).

Fig. 3 a, *a1* Transcription after release from the α -factor block. Purified G_0 daughters of strain M28 were shifted to 23 °C and grown out in YEPD for 6 h in the presence of α -factor, then inoculated into fresh YEPD in the presence or absence of α -factor. The levels of *a1* RNA are plotted against time after release. The proportion of budded cells in the release condition is also shown. **b, Effect of hydroxyurea (HU) or MBC.** The cells were released from α -factor as described in **a**, and inoculated into YEPD containing hydroxyurea or MBC. The levels of *a1* RNA are plotted against time after release. The proportion of budded cells in the MBC condition is also shown. In order to demonstrate that the blocks were effective, the cells were examined microscopically. Cells blocked at mitosis or at DNA replication arrest as cells with only one bud. A time-lapse microscopic analysis showed that <3 % of the cells produced a second bud in either the hydroxyurea or the MBC block. Furthermore, the levels of *HO* and histone H2B transcripts were also measured. As daughter cells were used, there should be no transcription of the *HO* gene in the first cell cycle, but if cells enter G_1 of the second cell cycle, *HO* transcription will begin³⁶. H2B is transcribed in late G_1 in all cells³⁸. Thus cells in α -factor should produce neither transcript, while those released into no cell cycle block at all should show considerable transcription of both *HO* and H2B. Cells released into hydroxyurea or MBC should show H2B transcription, but negligible *HO* transcription. This is what we observed (data not shown), demonstrating that both these blocks are effective. *a1* Transcription in the strain M29 was not affected by prolonged incubation in hydroxyurea. The pretreatment with α -factor minimizes the period that the cells are exposed to hydroxyurea or MBC, thus making any effect of growth inhibition by these reagents unimportant and also ensuring that the cells do not become 'habituated' to the poison and thus escape quite quickly from its effect²⁰. This procedure ensures that the block is tight and minimizes its intrusiveness. Daughter cells were used to maximize the overall growth at the permissive temperature (and the chance to synthesize enough functional *SIR3* protein) before any particular stage of the cell cycle was reached. Cells grown into G_0 at the permissive temperature and subjected to identical treatments did not show any significant silent copy transcription.

Methods: The daughters were made as described in Fig. 1 legend, then inoculated into YEPD to an A_{660} of 0.25; after 1 h at 23 °C, α -factor was added to 3 units ml^{-1} , and the cells were grown for a further 5 h, then collected by centrifugation and inoculated into YEPD containing 3 units ml^{-1} α -factor, or 200 mM hydroxyurea, or 50 $\mu g\ ml^{-1}$ MBC (1/200 volume of 10 $mg\ ml^{-1}$ MBC in dimethyl sulphoxide), or no cell cycle block at all. Cell samples for RNA preparation were chilled by the addition of about one-third volume of ice, total yeast RNA was prepared and levels of RNA were determined by S_1 nuclease protection of radiolabelled DNA, as described in ref. 36. The S_1 autoradiographs were analysed on a Transdyne 2955 gel scanner, and the ratio of the 245-nucleotide signal to the *SIR3* signal is plotted here. Since the multiple bands seen for the *SIR3* signal correspond to multiple termination points, or to S_1 nuclease 'chewing back', the sum of all of them was used. Initiation of the *SIR3* transcript was also measured and was not affected by these treatments.

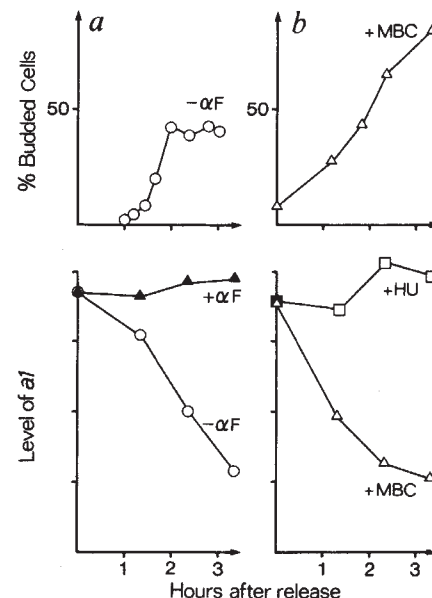
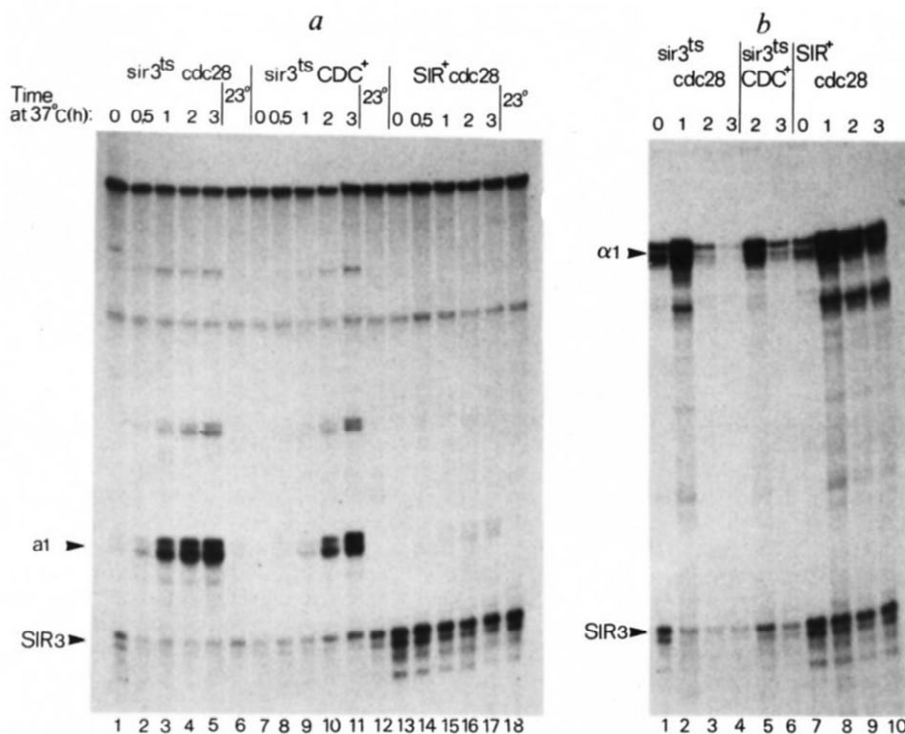


Fig. 4 a, The derepression of *HMRA* without DNA replication. Levels of *a1* RNA in a *MAT α sir^{ts}* strain after a shift to the restrictive temperature. Strains M20, M10 and M26 were grown into G_0 at the permissive temperature (23 °C) on YEPD plates and inoculated into fresh YEPD medium at the restrictive temperature (37 °C). Lanes 1–5, M20 (*sir^{ts} cdc28^{ts}*) after 0, 0.5, 1, 2 and 3 h at the restrictive temperature; lane 6, M20 grown out at the permissive temperature. Lanes 7–11, M10 (*sir^{ts} CDC⁺*) after 0, 0.5, 1, 2 and 3 h at the restrictive temperature. Lane 12, M10 grown out at the permissive temperature. Lanes 13–17, M26 (*SIR⁺ cdc28^{ts}*) after 0, 0.5, 1, 2 and 3 h at the restrictive temperature; lane 18, M26 grown out at the permissive temperature. The low level of transcript seen here could be due to the loss of the *SIR⁺* gene from its site of integration at the *TRP* locus. The effectiveness of the *cdc* block was confirmed by time-lapse microscopic analysis; in the *cdc28^{ts}* strains, less than 6 % of the cells produced buds. The difference in time course of derepression between M20 and M10 is due to background variation



between nonisogenic strains: M9 (Table 1) gives a time course identical to M10, and when M28 is subjected to the same temperature shift in the presence or absence of α -factor, both conditions gave full derepression after only 30 min (data not shown). **b, The repression of *MAT α 1* without DNA replication.** Levels of *a1* RNA in a *MAT α sir^{ts}* strain after a shift to the restrictive temperature. Some of the RNA samples from the experiment described in **a** were tested for levels of *a1* RNA by S_1 protection. Lanes 1–4, M20 after 0, 1, 2 and 3 h at the restrictive temperature. Lanes 5, 6, M10 after 2 and 3 h respectively at the restrictive temperature. Lanes 7–10, M26 after 0, 1, 2 and 3 h at the restrictive temperature. Levels of *a1* RNA were determined as described in Fig. 1 legend. 25 μg total RNA were used per reaction. The *a1* DNA probe was made by subcloning the *Xho* linker mutation, α X109 (ref. 39). This DNA was cut with *Xho*I and *Hinc*II and ligated to *Sal*I/*Sma*I-cut M13mp9. The identity of the recombinant plasmid was confirmed by chain termination sequencing⁴⁰.

the presence or absence of α -factor (data not shown). We conclude that cell cycle progress is not necessary for derepression and that *SIR3* activity is required in early G_1 to maintain repression. *SIR3*, therefore, does not just act during DNA replication.

The above data imply that *SIR* can act to repress silent mating type loci only on passage through S phase, whereas they may become derepressed without cell cycle progress. It is possible that an S phase requirement for repression is a general phenomenon and not peculiar to the mode of action of *SIR* regulation. We therefore tested whether there exist genes which can be repressed in the absence of DNA replication. The transcription of *MATa1* is repressed in diploids by the concerted action of the *MATa2* and *MATa1* gene products². $\alpha 1$ transcription is also repressed in a *sir⁻* cell expressing $\alpha 2$ and $\alpha 1$. A *MATa sir⁻* strain (such as M20) therefore transcribes $\alpha 1$ when grown to G_0 at the permissive temperature. We have just shown that $\alpha 1$ transcription from *HMRa* commences rapidly on inoculation into fresh medium at the restrictive temperature (even when cells are blocked in G_1). Repression of $\alpha 1$ should then occur. To determine whether such repression is dependent on cell cycle progress, we analysed the RNA samples from the previous experiment with an $\alpha 1$ DNA probe (Fig. 4b). Both M10 (*CDC⁺*) and M20 (*cdc28^{ts}*) show a transient increase in $\alpha 1$ transcription followed by a rapid reduction on inoculation at the restrictive temperature. We conclude that $\alpha 1$ transcription can be repressed by $\alpha 1/\alpha 2$ control in the absence of any progress through the cell cycle.

Discussion

We have shown that if *HML* or *HMR* is derepressed by growing a *sir3^{ts}* or a *sir4^{ts}* strain at its restrictive temperature, then the ensuing repression during growth at the permissive temperature can occur only on passage through S phase (Figs 1, 3). It seems likely that DNA replication is the critical event for *SIR* repression, although we cannot at present exclude the involvement of some other event in S phase. Our conclusions are based on the observation that repression can occur in the presence of MBC, which prevents mitosis (but not DNA replication), but not in the presence of α -factor or hydroxyurea, which prevent entry into S phase. Unfortunately, the use of a *ts sir⁻* allele to derepress *HML* and *HMR* has precluded the use of *ts* DNA replication mutants to pursue this question any further.

The requirement for S phase is significant, however, because the *cis*-acting control sequences (E_L and E_R), via which *SIR* represses *HML* and *HMR*, have ARS activity¹². We propose that DNA replication initiated at E is essential for repression. We believe that an investigation of the role of the ARS activity of E is the most promising approach for establishing the involvement of DNA replication in *SIR* repression.

The mechanism of *SIR* control can be formally divided into three steps. First, a *SIR* protein recognizes and, presumably, binds at E. Second, this signal is translocated to the target promoters, and third, the transcription of these promoters is repressed. In previously characterized systems of negative control (for example, *lac²⁴* or bacteriophage λ^{25}), these three steps (recognition, translocation and repression) are coincident; the binding of the repressor to the operator seems sufficient for repression and there is no need to translocate any signal. *SIR* control appears to be different since the site of repression may be many hundreds of base pairs away from the recognition site (E), just as the recognition site for N-antitermination is distinct from the actual site where antitermination occurs^{26,27}.

DNA replication may be required for the recognition of E, or for the act of repression, or for communication between the two. For example, the functional *SIR3* (or *SIR4*) protein produced during growth at the permissive temperature may not be able to bind to E unless it is opened up at replication. In this instance, binding during replication may be a means of forming a complex that cannot be dissociated until the next round of replication. Alternatively, DNA replication may only be required for the act of repression: for instance, the *HMLa1* and

HMRa1 promoters may bind a transcription factor which cannot be removed from the DNA except at replication. Thus, once the silent copies have been derepressed, this factor is bound to the promoter and the *SIR⁺* or *sir⁻* status of the cell is irrelevant until DNA replication occurs. Similarly, repression may be achieved by the construction of some repressed chromatin structure which can only be produced as the nascent DNA is packaged in S phase²⁸. Since derepression can occur without progress through the cell cycle (Fig. 4a), however, this repressed state must be dependent on the continued action of the *SIR3* gene product. The latter point would seem to exclude replication-dependent methylation as the mode of *SIR* repression. Another possibility is that the translocation of the signal is replication-dependent, as would be the case if translocation were achieved by the movement of a replication fork initiated at E to the affected promoters.

Whichever of these different hypotheses is correct, our results suggest that some modulator of transcription is unable either to exchange onto, or off, the DNA in the absence of DNA replication.

Gene expression is linked to DNA replication in both prokaryotic and eukaryotic viruses. The early-late switch in transcription accompanies, and appears to be dependent on, viral DNA replication. There are often important biological reasons for this pre-replication/post-replication distinction in viral transcription. In some cases, the apparent 'requirement' for DNA replication may be indirect, and have little to do with the actual mechanism of transcriptional control. For example, it has been proposed that DNA replication is required for late gene expression in SV40 (ref. 29). More recent work, however, suggests that the late promoter can be activated by T-antigen in the absence of DNA replication³⁰. In a bacteriophage T_4 infection, late transcription requires some form of template activation, but continuing DNA replication appears to be just one way of achieving this^{31,32}. In adenovirus, on the other hand, it has been demonstrated that a replicated DNA template is transcribed differently from an unreplicated DNA template within the same cell³³, suggesting that DNA replication may be involved directly in the mechanism of the early-late switch.

Unlike viral systems, where the coordination of gene expression and DNA replication is part of a developmental pathway, *SIR* control in yeast is normally continuous throughout the cell cycle of either haploid or diploid cells. Our finding that replication is required for establishing repression (following derepression using a *ts sir* mutant) need not suggest therefore that *SIR* control is developmentally regulated, but more probably reflects the actual mechanics of *SIR* repression.

A replication-dependent mechanism of transcriptional control could be important in the maintenance of developmental switches in gene expression. Whereas the signals and circumstances inducing cell differentiation are often transient, the subsequent patterns of gene expression in the terminally differentiated cell are often permanent. One mechanism by which this can be achieved is the existence of positive feedback loops, as in bacteriophage λ^{25} . An alternative mechanism could involve the final division of a differentiating cell leading to a state of gene expression which cannot be altered in the absence of further DNA replication. For instance, our findings suggest that yeast cells will continue to transcribe the *HMRa1* gene despite the presence of *SIR* repressors as long as the cells are prevented from undergoing DNA replication.

We thank Rolf Sternglanz, Linda Breeden, David Shore and Alfonso Martinez-Arias for helpful comments.

Received 9 May; accepted 21 August 1984.

1. Astell, C. R. *et al. Cell* **27**, 15–23 (1981).
2. Nasmyth, K. A., Tatchell, K., Hall, B. D., Astell, C. R. & Smith, M. *Nature* **289**, 244–250 (1981).
3. Hawthorne, D. C. *Abstr. Proc. 11th int. Congr. Genet.* **7**, 34–35 (1963).
4. Herskowitz, I. & Oshima, Y. in *The Yeast Saccharomyces cerevisiae* (eds Strathern, J. N., Jones, E. W. & Broach, J.) 181–211 (Cold Spring Harbor Laboratory, New York, 1981).
5. Klar, A. J. S., Strathern, J. N., Broach, J. B. & Hicks, J. B. *Nature* **289**, 239–244 (1981).
6. Abraham, J., *et al. Cold Spring Harb. Symp. quant. Biol.* **47**, 989–997 (1983).
7. Haber, J. E. & George, J. P. *Genetics* **93**, 13–35 (1979).

8. Klar, A. J. S., Fogel, S. & Macleod, K. *Genetics* **93**, 37-50 (1979).
9. Rine, J., Strathern, J. N., Hicks, J. B., & Herskowitz, I. *Genetics* **93**, 877-901 (1979).
10. Rine, J. O., thesis, Univ. Oregon (1979).
11. Banerji, T., Rusconi, S., & Schaffner, W. *Cell* **27**, 299-308 (1981).
12. Abraham, J., Nasmyth, K. A., Strathern, J. N., Klar, A. J. S. & Hicks, J. B. *J. molec. Biol.* **176**, 307-331 (1984).
13. Stinchcomb, D. T., Struhl, K. & Davies, R. W. *Nature* **282**, 39-43 (1979).
14. Beach, D., Piper, M. & Shall, S. *Nature* **284**, 185-187 (1980).
15. Chan, C. S. M. & Tye, B.-K. *Proc. natn. Acad. Sci. U.S.A.* **77**, 6329-6333 (1980).
16. Newlon, C. S. & Burke, W. in *Mechanistic Studies of DNA Replication and Genetic Recombination* Vol. 19 (eds Alberts, B. & Fox, C. F.) 339-409 (Academic, New York, 1980).
17. Celniker, S. E. & Campbell, J. L. *Cell* **31**, 201-213 (1982).
18. Miller, A. M. *EMBO J.* **3**, 1061-1065 (1984).
19. Thorne, J. in *The Yeast Saccharomyces cerevisiae* (eds Strathern, J. N., Jones, E. W. & Broach, J.) 143-180 (Cold Spring Harbor Laboratory, New York, 1981).
20. Hartwell, L. H. *J. molec. Biol.* **104**, 803-817 (1976).
21. Kilnlan, R. A., Pogson, C. I. & Gull, K. J. *Cell* **46**, 341-352 (1980).
22. Kilmartin, J. V. *Biochemistry* **20**, 3629-3633 (1981).
23. Hereford, L. & Hartwell, L. H., *J. molec. Biol.* **84**, 445-461 (1974).
24. Gilbert, W. & Muller-Hill, B. in *The Lactose Operon* (eds Beckwith, J. R. & Zipser, D.) 93-109 (Cold Spring Harbor Laboratory, New York, 1970).
25. Johnson, A. D. *et al. Nature* **294**, 217-223 (1981).
26. Salstrom, J. S., & Szybalski, W. *J. molec. Biol.* **124**, 195-221 (1978).
27. de Crombrughe, B., Mudryj, M., DiLauro, R. & Gottesman, M. *Cell* **18**, 1145-1151 (1979).
28. Alberts, B., Worcel, A. & Weintraub, H. in *The Organization and Expression of the Eukaryotic Genome* (eds Bradbury, E. M. & Javaherian, K.) 165-191 (Academic, New York, 1977).
29. Contreras, R., Gheysen, D., Knowland, J., van de Voode, A. & Fiers, W. *Nature* **300**, 500-505 (1982).
30. Keller, J. M. & Alwine, J. C. *Cell* **36**, 381-389 (1984).
31. Geiduschek, E. P., Elliott, T. & Kassavetis, G. A. in *Bacteriophage T4* (eds Mathews, C. K., Kutter, E. M., Mosig, G. & Berget, P. B.) 189-192 (Am. Soc. Microbiol., Washington, DC, 1980).
32. Elliott, T. & Geiduschek, E. P. *Cell* **36**, 211-219 (1984).
33. Thomas, G. P., & Mathews, M. B. *Cell* **22**, 523-533 (1980).
34. Hartwell, L. J. *Cell Biol.* **85**, 811-822 (1980).
35. Reed, S. I. *Genetics* **95**, 561-577 (1980).
36. Nasmyth, K. A. *Nature* **302**, 670-676 (1983).
37. Bucking-Throm, E., Duntze, W., Hartwell, L. H. & Manney, T. R. *Expl Cell Res.* **76**, 99-110 (1973).
38. Hereford, L., Bromley, S. & Osley, M. A. *Cell* **30**, 305-310 (1982).
39. Tatchell, K., Nasmyth, K. A., Hall, B. D., Astell, C. R. & Smith, M. *Cell* **27**, 25-35 (1981).
40. Sanger, F., Nicklen, S. & Coulson, A. R. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5463-5467 (1977).

LETTERS TO NATURE

Dynamic spectra of interplanetary scintillations

W. A. Coles & J. P. Filice

Department of Electrical Engineering and Computer Sciences,
University of California, San Diego, La Jolla, California 92093, USA

Interplanetary scintillations (IPS) are fluctuations in the apparent intensity of radio sources caused by fluctuations in the electron density of the solar wind¹. Observations at a single wavelength provide information on the spatial spectrum of electron density at scales of the order of the radius of the first Fresnel zone. However, dynamic spectra covering a wide range of wavelengths show evidence of fluctuations on a much larger scale². Unfortunately, the problem of estimation of the spatial spectrum at this larger scale is analytically intractable. Here we compare the observations² with a numerical simulation of the diffraction of plane waves by a two-dimensional random phase screen. The frequency-time structure of the simulated IPS shows good qualitative agreement with the observations when the turbulence spectrum in the phase screen is a simple power law with the Kolmogorov exponent. No *ad hoc* large-scale structure is necessary to account for the observations.

Observations of dynamic spectra over a wide range require an unusual spectrometer and have been reported only by Cole and Slee²; they observed IPS of the source 3C273 using an acousto-optical spectrograph on the Parkes telescope to cover a band of 280-520 MHz. At the time of these observations (10-11 October 1978), the solar elongation was 13°, so that the IPS were strong at the lower frequencies and weak at the higher frequencies. Figure 1 gives a sample of these data. Bands of

enhanced intensity are evident, curving in both senses with frequency. The bands take about 1 s to sweep across the frequency range.

Cole and Slee² suggested that these bands are due to "systematic refraction from sharp density gradients", and Hewish, in an addendum to their paper, advanced the alternative explanation that refraction from large-scale components of a normal turbulence spectrum was responsible. A similar phenomenon has been observed in the interstellar scintillation (ISS) of pulsars^{3,4}. Although curvature is not evident in ISS, presumably because the bandwidths are smaller, tilted bands are clearly observed. Shishov⁵ accounted for this by postulating two characteristic scales in the medium: a small scale responsible for the scintillation, and a large scale (too large to produce appreciable intensity fluctuations at the Earth) which produces a refractive tilt to the wavefront, shifting the small-scale pattern laterally, alternately squeezing and stretching it. Rickett and Lang⁶ proposed a similar explanation to account for the apparent reversals in the drift velocity of ISS patterns observed with two space antennas. As these refractive shifts depend systematically on frequency, they can explain the frequency-time drifts observed.

Clearly, the tilt of the bands is caused by larger structures than those responsible for the intensity fluctuations themselves. Shishov⁷ showed that such observations could be used to estimate the variance in these large-scale structures. Gapper and Hewish⁸ recently completed a programme of two-frequency IPS observations intended to do just that; these observations were found to be consistent with a power-law spectrum of the Kolmogorov exponent. This work is limited, however, by the fact that the observations exist at only two frequencies and by the absence of an exact theory. Thus, it is difficult to make more than a consistency argument.

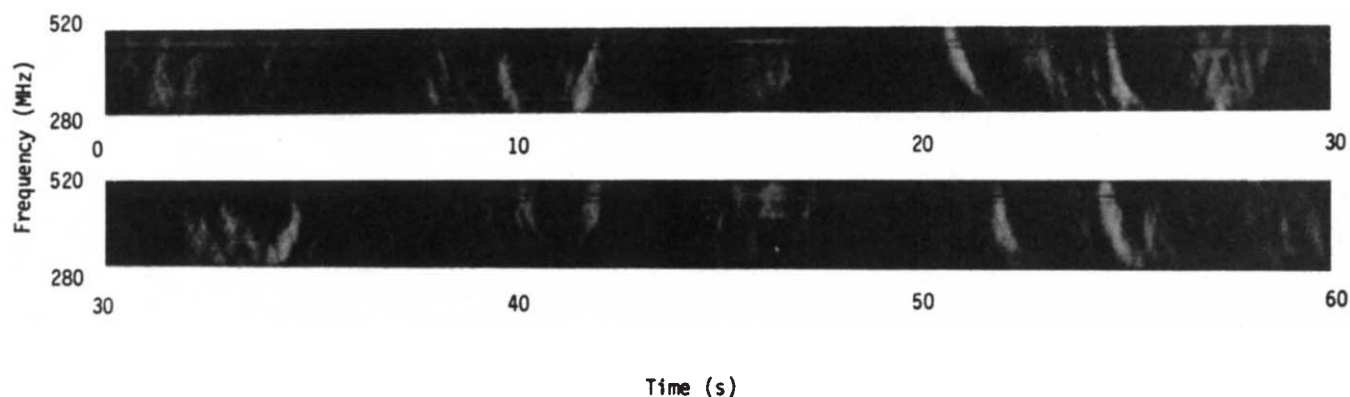


Fig. 1 An example of the intensity variations of the source 3C273 due to interplanetary scintillation. Systematic structure on the frequency-time plane is evident. The figure is plotted with a frequency resolution of 4 MHz, time resolution of 40 ms, and white represents increased intensity (Fig. 1 of Cole and Slee²).

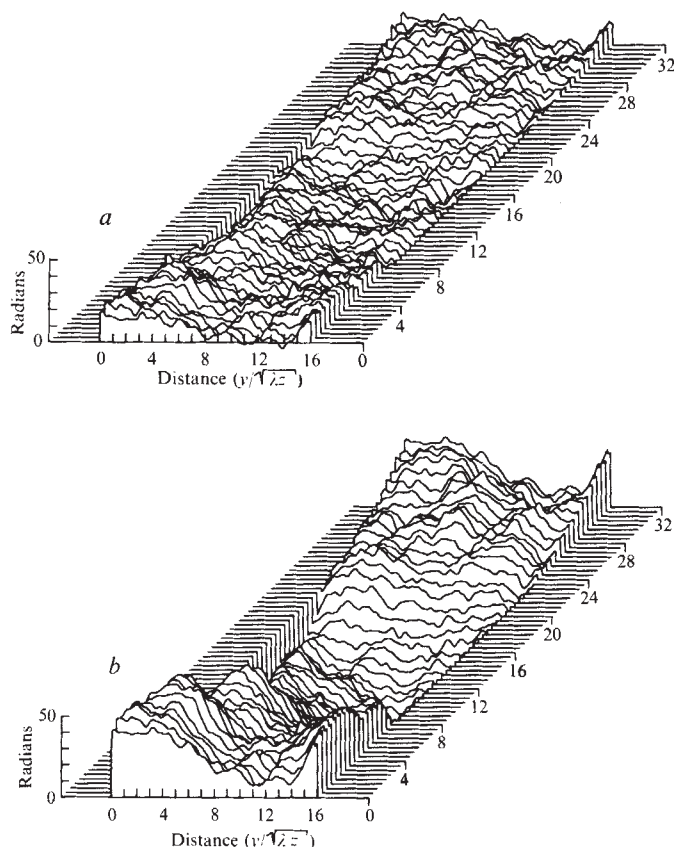


Fig. 2 Two random-phase screens computed for $U = 10$ from the same pseudorandom number sequence. The specified phase power spectrum is power law in each case with exponents: $\alpha = 3$ (a); and $\alpha = 11/3$ (b). There are 512 points in the long direction and 256 points in the short direction; however, only every eighth cut through the short dimension is shown. Each cut has been smoothed with a running mean 8 samples wide.

The frequency-time structure of IPS is a difficult phenomenon to analyse. The interesting structure evident in the Cole and Slee observations² occurs in the region of saturated scintillations. The moment equations which describe narrow forward angular scatter of radio waves by random media are intractable in this region. In principle, the moment equations could be solved numerically to gain insight into the problem but at present there are insufficient observations of the process from which to extract meaningful statistics. Furthermore, it is not clear which moment of the electric field is most relevant. Therefore, we have chosen to approach the problem by performing simulations in which realizations of the random process are generated digitally.

A simulation has certain advantages over a purely analytical or numerical approach. Each realization of the process is generated from a sequence of pseudorandom numbers which is scaled by the turbulence spectrum. Therefore, the effect on a particular realization produced by a change in the spectrum can be obtained directly. Also, the scattered field from each realization can be compared with the scattering 'medium' in a detailed way which is not possible with actual scintillation observations. Both of these features, peculiar to simulations, are exploited in the present study to explore the effect of large-scale phase gradients on IPS.

At an elongation of 13° , it is accurate to treat the scattering medium as if it were concentrated into a thin phase-changing layer at the closest point on the line of sight to the Sun. The random phase is needed over a two-dimensional grid of finite extent; thus it has a discrete Fourier transform of the form $\tilde{\phi}_{kl} = (A_{kl} + iB_{kl})/\sqrt{2}$ where A_{kl} and B_{kl} are independent, zero-mean gaussian random variables and k and l are the indices of the two-dimensional grid. The variance σ_{kl}^2 of A_{kl} and B_{kl} is

related to the phase power spectrum by $\sigma_{kl}^2 = L_x L_y \Phi_p(\kappa_{kl})$ where L_x and L_y are the dimensions of the phase screen and κ_{kl} is the two-dimensional wavenumber. Thus we have used a pseudorandom sequence, scaled by $\Phi_p(\kappa_{kl})$, to generate A_{kl} and B_{kl} . The desired phase screen is then obtained by discrete Fourier transformation.

The phase spectrum used was of the form $\Phi_p = 2\pi T \kappa^{-\alpha}$ where T is a real constant. The spectral exponent α of this two-dimensional phase spectrum is the same as that of the three-dimensional electron density spectrum. Although we have limited ourselves here to the simple isotropic power law form, anisotropy and complex structure can easily be introduced into the spectrum.

It is convenient to express the 'strength of turbulence' in terms of a parameter U , which is exactly the Born (weak scintillation) approximation to the intensity variance⁹. This is a useful 'strength' parameter for use with intensity scintillations as $U < 1$ in weak scintillations and $U > 1$ in strong scintillations. In effect U is an integral over that portion of the phase spectrum near the Fresnel wavenumber $\kappa_f = \sqrt{2\pi/\lambda z}$ where z is the distance from screen to observer and λ is the wavelength of the incident plane wave. This relation between U and T is given by

$$U = -T\Gamma\left(1 - \frac{\alpha}{2}\right) \sin\left(\frac{\alpha\pi}{4}\right) \left(\frac{\lambda z}{2\pi}\right)^{\alpha/2-1} \quad 2 < \alpha < 6, \alpha \neq 4,$$

$$U = \frac{T\lambda z}{4} \quad \alpha = 4$$

The diffraction pattern produced by a plane wave incident on the screen was calculated using an integral solution to the parabolic wave equation¹⁰ which describes narrow-angle forward scattering. Details of the phase screen generation and the computation of the diffraction pattern are given by Filice¹¹.

Because the full complex electric field is available, any moment of the field may be calculated. The ensemble average of any statistic of the field is estimated by averaging many realizations. This provides a convenient check of the simulation in two ways: a simple and exact formula is available for the spatial correlation function of the electric field, and the power spectrum of intensity can be calculated easily in weak scattering. The number of samples required to adequately represent the two-dimensional phase screen increases as U^4 when the scintillations are strong. Data storage and calculation time constraints limit the number of samples with which to represent the screen. Thus, there is a trade-off between the 'finesses' of a screen and its spatial extent. These choices were tested by comparing the field computed from a phase screen with the field computed from the same phase screen after it had been windowed and/or decimated. Finally, the second-order statistics of the electric field and the intensity were compared, where possible, with the theory.

Phase screens generated from the same random number sequence for $\alpha = 3$ and for $\alpha = 11/3$ are shown in Fig. 2. These screens consist of a 512×256 array of phase shifts; however, for display purposes, only every eighth cut through the screen is shown and each cut has been smoothed using a running mean 8 samples wide.

Dynamic spectra computed using the two screens are shown in Fig. 3; they were generated by computing the diffraction pattern produced by the screen at 23 frequencies. Only a cut through the centre of the diffraction pattern is used at each frequency. Curved bands of enhanced intensity are evident. These examples are typical of several dozen calculated and in each case bands of enhanced intensity are evident. The bands appear to break up for $U \geq 5$ in the case $\alpha = 3$ but for $\alpha = 11/3$ they remain well defined over the entire range of U . The curvature of the bands is much greater for $\alpha = 11/3$ than for $\alpha = 3$.

The observations show variations of intensity which are highly correlated over the frequency range and take ~ 1 s to sweep across the frequency range. If the solar wind velocity is 400 km s^{-1} , each simulated dynamic spectrum corresponds to

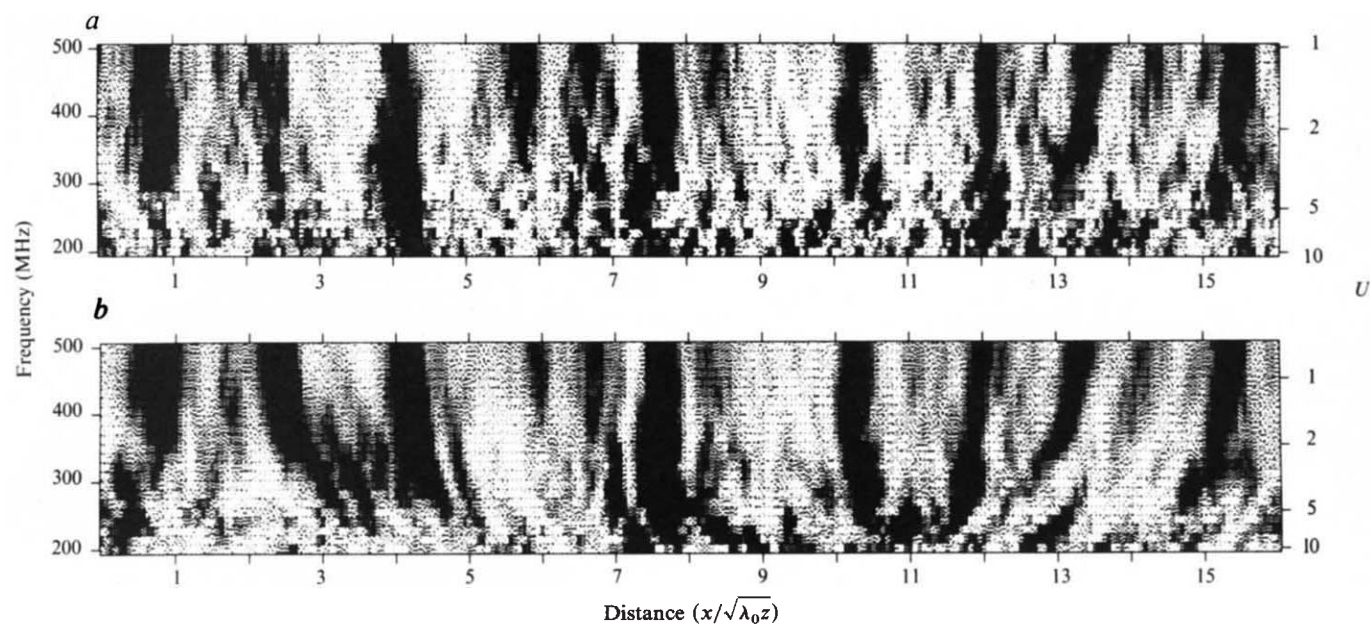


Fig. 3 Simulated dynamic spectra computed for *a*, $\alpha = 3$ and *b*, $\alpha = 11/3$ from the phase screens of Fig. 2. Black represents increased intensity and saturates at three times the mean level. The horizontal distance is normalized by $\sqrt{\lambda_0 z}$ where $\lambda_0 = 1.5$ m is the wavelength corresponding to the lowest frequency observed (strongest scattering) and z is the distance from observer to screen. For convenience, the corresponding intensity randomization is shown at the righthand side of each dynamic spectrum. The conversion from frequency f to U is $U = (f_0/f)^{1+\alpha/2} U_0$ where f is in MHz. Here $f_0 = 200$ MHz and $U_0 = 10$.

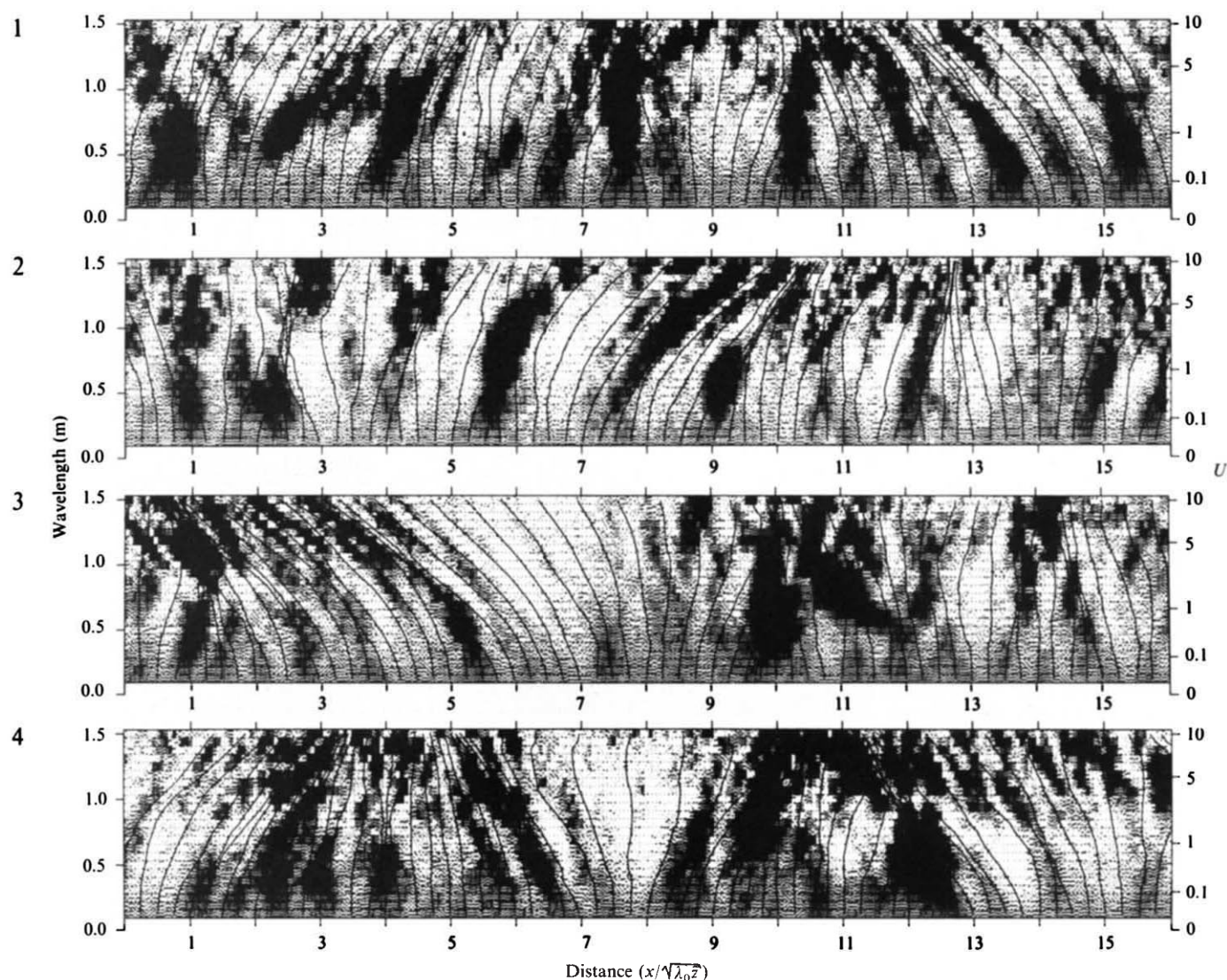


Fig. 4 Four realizations of dynamic spectra with $\alpha = 11/3$, showing the predicted shift of intensity features due to refraction by large-scale phase gradients.

20 s of time. The bands of correlated intensity in the simulated dynamic spectra with $\alpha = 11/3$ have slopes of the same order.

Although qualitatively there is good agreement between the simulation and the observations, it is difficult to establish the degree of quantitative agreement. It is possible to estimate the range of U observed by Cole and Slee². It is known from IPS observations at 74 MHz¹² and the numerical calculations of Marians¹³ that, on average, $U \sim 35$ at an elongation of 13° . As $U \propto \nu^{-(1+\alpha/2)}$ then at $\nu = 280$ MHz, $U \sim 1.3$ and at $\nu = 520$ MHz, $U \sim 0.3$ and with reasonable certainty $0.1 < U_{\text{obs}} < 3$.

Because the scattered intensity can be compared directly with the phase screen, we can test the hypothesis that large-scale phase gradients are responsible for the shift of the intensity pattern with frequency.

In the small-angle approximation, the angular tilt of a wavefront due to a phase gradient is

$$\delta_d = \frac{\lambda}{2\pi} \nabla \phi(x, y)$$

and the lateral deviation at distance z is $d = z\delta_d$. Since for a plasma $\phi \propto \lambda$ then $d \propto \lambda^2$. The shift of the pattern is due to the component of $\nabla \phi$ parallel to the velocity or, in this case, $\partial \phi / \partial x$.

To test the phase gradient hypothesis, it is assumed that a particular field point is due to scattering from a disk on the phase screen of radius $\theta_s z$, where θ_s is the width of the angular spectrum of plane waves in the scattered wave. For the Kolmogorov spectrum

$$\theta_s \sim \frac{(3U)^{3/5}}{2\pi} \left(\frac{\lambda}{z} \right)^{1/2}$$

At equi-spaced positions along the screen, the deviation of the small-scale pattern was estimated by (1) fitting the region of the screen defined by θ_s with a plane, (2) calculating $\partial \phi / \partial x$, and (3) computing the lateral shift of the pattern from the expression above. This calculation was performed at each of the 23 wavelengths used in the simulation. These shifts are plotted as tracks in the simulated dynamic spectra shown in Fig. 4. The spectra have been computed with steps of equal wavelength to expand the strong scattering region where most of the curvature occurs. The undeviated positions are seen at the bottom of the dynamic spectra where the scattering is weak and the deviation small. The plotted lines show how the intensity pattern produced by small-scale phase fluctuations would shift due to large-scale fluctuations. The predictions follow the intensity bands quite closely, lending significant weight to this simple argument.

The range of spatial scales in the solar wind which contribute to a point observation range from the coherence scale of the electric field, s_0 , to the size of the 'scattering disk', s_L . For $U = 3$ at $\nu = 280$ MHz, $s_0 \sim 70$ km and $s_L \sim 500$ km. However, the curved bands evident in the dynamic spectra show correlation over still larger scales in the sense that the gradient changes sign on a scale of $\sim 2,000$ km. This comparison indicates, therefore, that the mean spectral exponent over this range (at least 70–2,000 km) is comparable with the Kolmogorov value of $-11/3$. An exponent of -3 would be too small to produce the observed curved bands of enhanced intensity.

Solar neutrinos and other problems and their relation to energy production in the Sun

C. L. Joseph

Department of Astronomy, The Ohio State University, Columbus, Ohio 43210, USA

Models of the solar interior, based on the usual physical assumptions¹, predict a neutrino flux several times greater than that observed in the Davis ³⁷Cl experiment^{2,3}. If, as is widely accepted, this discrepancy represents a 'flaw' in the standard solar model one would expect this flaw to manifest itself in other ways also. Here we point out some less well known discrepancies between theoretical predictions of the standard solar model and relevant observations. We suggest that an additional source of energy production that competes with the proton-proton chain and carbon-nitrogen cycles, can reduce these discrepancies if it has a sufficiently strong temperature dependence. Model calculations based on a possible new mode of nucleosynthesis^{4,5} support this conclusion.

The identification of the 5-min solar oscillations with acoustical normal modes trapped beneath the solar photosphere^{6–8} has provided a powerful probe of the solar interior. Recently, Ulrich and Rhodes⁹ showed that the discrepancy between the observed and theoretically predicted frequencies of the low-order modes cannot be accounted for by either errors in the observations themselves or their treatment and thus represents a real failure of the standard solar model. Specifically, the observations suggest that the solar convection zone is considerably deeper than predicted by the standard models^{10–12}. A deeper convective envelope is also indicated by the low ⁷Li abundance inferred from photospheric spectra^{13,14} and the absence of rapid rotation near the solar poles¹⁵.

Another prediction of the standard model is the fractional helium abundance by mass (Y) of the zero-age Sun. Most solar models^{16–18} predict a helium abundance in the range $0.21 < Y < 0.25$, which differs little from the primordial value predicted from standard Big Bang cosmology^{19–21} and inferred from observations of metal-deficient interstellar HII regions²². It is difficult to reconcile such a low solar helium abundance with our present understanding of galactic chemical evolution. Indeed, solar chromospheric spectra²³ suggest $Y = 0.28$, which is consistent with that found in other objects having almost the solar abundance of heavy elements^{24,25}.

Consider the effect on the solar structure of an additional, centrally concentrated energy source characterized by a strong temperature dependence. Because of this energy source the star is more nearly described by a Cowling point-source model²⁶. The increased efficiency of the energy sources in such a model results in a lower luminosity, a lower central mass concentration and hence a larger radius than in the standard solar model²⁶ (assuming that the time dependence of the luminosity and radius is not appreciably altered). As a model of the present Sun is constrained to have the observed solar radius and luminosity, the addition of this energy source requires that we adjust one or more physical parameters in order to satisfy these constraints. We can obtain the present solar luminosity, at an adopted solar age of 4.7×10^9 yr, by increasing the initial helium abundance (through the dependence of the luminosity on the mean molecular weight and opacity). The present solar radius can be recovered by increasing the value of the mixing length l , which determines the adiabat deep in the convection zone. A larger value of l corresponds to more efficient convection in the outer layers, a flatter temperature gradient and therefore a higher surface temperature. Because a change in l has little effect on the luminosity, a higher surface temperature results in a smaller radius. Consequently, the depth of the solar convection zone must be increased (via an increase in l) in order to match the solar radius. The necessity for a deeper convective envelope is

Received 18 June; accepted 23 August 1984.

- Hewish, A., Scott, P. F. & Wills, D. *Nature* **203**, 1214–1217 (1964).
- Cole, T. W. & Slee, O. B. *Nature* **285**, 93–95 (1980).
- Ewing, M. S., Batchelor, R. A., Friefeld, R. D., Price, R. M. & Staelin, D. H. *Astrophys. J.* **162**, L169–L172 (1970).
- Roberts, J. A. & Ables, J. G. *Mon. Not. R. astr. Soc.* **201**, 1119–1138 (1982).
- Shishov, V. I. *Soviet Astr.* **17**, 598–602 (1974).
- Rickett, B. J. & Lang, K. R. *Astrophys. J.* **185**, 945–950 (1973).
- Shishov, V. I. *Soviet Astr.* **19**, 226–228 (1975).
- Gapper, G. R. & Hewish, A. *Mon. Not. R. astr. Soc.* **197**, 209–216 (1981).
- Rumsey, V. H. *Radio Sci.* **10**, 107–114 (1975).
- Tatarskii, V. I. *The Effects of the Turbulent Atmosphere on Wave Propagation* (transl. by Israel Program for Scientific Translations) (National Technical Information Service, U.S. Dept of Commerce, Washington, DC, 1971).
- Filice, J. P. thesis, Univ. California, San Diego (1984).
- Rickett, B. J. & Seiradakis, J. H. *Astrophys. J.* **256**, 612–616 (1982).
- Marians, M. *Radio Sci.* **10**, 115–119 (1975).

also a natural consequence of the steeper central temperature gradient; the radiative gradient intercepts the adiabatic gradient at a greater depth beneath the photosphere.

As the core of the standard solar model is only marginally stable against convection, it is likely that even a modest increase in the central temperature gradient would be sufficient to induce convection. If this were the case, one would expect the main sequence evolution of the Sun to be qualitatively similar to that of a slightly more massive star possessing a convectively mixed core. It is interesting, in this context, that the colour-magnitude diagrams of some old open clusters show an anomalous gap near the main sequence turnoff, usually attributed to core mixing of an unknown origin²⁷. The Sun's main sequence evolution would be towards the right in the Hertzsprung–Russell diagram, with a relatively smaller luminosity increase than a star with a radiative interior. If the luminosity of the early Sun were in fact larger than predicted by the standard models, this would help to explain the palaeoclimatic evidence which suggests that the global mean temperature was not markedly different in the past^{28,29}. Furthermore, the reduction of the change in past solar luminosity together with the increased initial helium abundance implies a shorter main sequence lifetime. This would lead to a reduction in the ages of globular clusters as inferred from standard stellar evolution theory. Comparison of globular cluster colour-magnitude diagrams with existing theoretical isochrones implies ages of the order of 18×10^9 yr for the oldest clusters^{30,31}, which is disturbingly near the upper limit imposed by the value of the Hubble constant³².

Possible candidates for the additional energy source have been discussed elsewhere^{4,5,33}. Boyd *et al.*^{4,5} recently proposed a mode of nuclear energy production based on the possible existence of nuclei containing an extra embedded particle, for example, an extra quark or an X-particle³⁴, which would increase the nucleon binding energy. The existence of such entities is not precluded by contemporary theories of particle physics; indeed, numerous experimental searches have been suggested and some are in progress^{4,34}. Assuming that the embedded particle is a quark, Boyd *et al.* show that a hydrogen-burning cycle catalysed by these 'Q-nuclei' could contribute to solar energy production without producing any detectable neutrinos. In this cycle, ^4He is produced by successive proton captures on He, Li and Be nuclei containing an extra quark. As the collision cross-section $\langle\sigma v\rangle$ for the slowest reaction in this cycle is many orders of magnitude larger than that of the proton-proton reaction at typical solar temperatures, even a minute abundance of Q-nuclei would be sufficient for this cycle to contribute significantly to the Sun's energy output. Furthermore, the Q-nuclear reaction rates fulfil the large temperature sensitivity requirement.

We have made a study of solar models incorporating these Q-nuclear reaction rates, using a modified version of the Paczynski evolution code³⁵. The principal revisions to the code involved upgrading the physics and improving the overall efficiency; the standard solar model obtained is in good agreement with that of other investigators^{3,16}, yielding a neutrino capture rate of 7 SNU (1 SNU = 10^{-36} captures per ^{37}Cl atom s^{-1}). For this investigation, the equations of nucleosynthesis were modified to include ^4He production via the Q-nuclear cycle, only the slowest reaction in the cycle being considered explicitly. As the Q-nuclei act as catalysts, their abundance (a free parameter) was assumed to be constant during the evolution. The predicted neutrino capture rate agrees with the observed value at an abundance of $\sim 10^{-15}$ Q-nuclei per normal nucleus. In this model, the depth of the outer convective zone is $\sim 20\%$ greater than that of the standard solar model, terminating at a temperature of $\sim 2.5 \times 10^6$ K. This temperature is within the range needed¹³ to account for the observed depletion of ^7Li in the solar photosphere. The initial helium abundance, Y , is increased by ~ 0.04 , in good agreement with other helium abundance determinations^{23–25}. The presence of a small ($\sim 0.05 M_\odot$) convective core results in a luminosity increase of only 28% over the past 4.7×10^9 yr, compared with 44% for standard models³.

As the energies of the neutrinos produced by Q-nucleosynthesis are expected to be somewhat different from those of the standard solar model⁵, solar neutrino detectors other than ^{37}Cl , for example ^{71}Ga , might provide a test for this mode of energy production. Mass spectroscopic searches for Q-nuclei may also prove fruitful. Furthermore, an eigenfrequency analysis may determine whether the structure of the proposed model is compatible with the observed acoustical spectrum of the Sun. Work is now underway to obtain more quantitative predictions from this spectrum and about other aspects of the solar structure.

We thank Drs George W. Collins II, Richard N. Boyd and Gerald H. Newsom for many helpful comments and suggestions.

Received 12 June; accepted 27 September 1984.

1. Bahcall, J. N. & Sears, R. L. *A. Rev. Astr. Astrophys.* **10**, 25–44 (1972).
2. Bahcall, J. N. *Space Sci. Rev.* **24**, 227–251 (1979).
3. Bahcall, J. N., Heubner, W. J., Lubow, S. H., Parker, P. D. & Ulrich, R. K. *Rev. mod. Phys.* **54**, 767–799 (1982).
4. Boyd, R. N., Turner, R. F., Wiescher, M. & Rybaczky, L. *Phys. Rev. Lett.* **51**, 609–612 (1983).
5. Boyd, R. N., Turner, R. E., Rybaczky, L. & Joseph, C. *Astrophys. J.* (in the press).
6. Ulrich, R. K. *Astrophys. J.* **162**, 993–1002 (1970).
7. Wolf, C. L. *Astrophys. J.* **176**, 833–842 (1972).
8. Ando, H. & Osaki, Y. *Publ. astr. Soc. Jap.* **27**, 581–585 (1975).
9. Ulrich, R. K. & Rhodes, E. J. *Astrophys. J.* **265**, 551–563 (1983).
10. Scuflaire, R., Gabriel, M. & Noels, A. *Astr. Astrophys.* **99**, 39–42 (1981).
11. Gabriel, M., Scuflaire, R. & Noels, A. *Astr. Astrophys.* **110**, 50–53 (1982).
12. Berthomieu, G. *et al. Lect. Notes Phys.* **125**, 307–312 (1980).
13. Weymann, R. & Sears, R. L. *Astrophys. J.* **142**, 174–181 (1965).
14. Boesgaard, A. M. *Publ. astr. Soc. Pacific* **88**, 353–366 (1976).
15. Gilman, P. A. *Astrophys. J.* **231**, 284–292 (1979).
16. Mazzitelli, I. *Astr. Astrophys.* **79**, 251–253 (1979).
17. Gehren, T. *Astr. Astrophys.* **109**, 187–190 (1982).
18. Sweigert, A. V. *Astrophys. J.* **273**, 795–804 (1983).
19. Schramm, D. N. & Wagoner, R. V. *A. Rev. nucl. Sci.* **27**, 37 (1977).
20. Yang, J., Schramm, D. N., Steigman, G. & Rood, R. T. *Astrophys. J.* **227**, 697–704 (1979).
21. Olive, K. A., Schramm, D. N., Steigman, G., Turner, M. S. & Yang, J. *Astrophys. J.* **246**, 557–568 (1981).
22. Kunth, D. & Sargent, W. L. W. *Astrophys. J.* **273**, 81–98 (1983).
23. Heasley, J. N. & Milkey, R. W. *Astrophys. J.* **221**, 677–688 (1978).
24. Peimbert, M. A. *Rev. Astr. Astrophys.* **13**, 113–131 (1975).
25. Peimbert, M. & Torres-Peimbert, S. *Mon. Not. R. astr. Soc.* **179**, 217–234 (1977).
26. Cox, J. P. & Giuli, R. T. *Principles of Stellar Structure*, 723–736 (Gordon & Breach, New York, 1968).
27. Maeder, A. *Astr. Astrophys.* **32**, 177–190 (1974).
28. Knauth, L. P. & Epstein, S. *Geochim. cosmochim. Acta* **40**, 1095–1108 (1976).
29. Newkirk, G. Jr. *A. Rev. Astr. Astrophys.* **21**, 429–467 (1983).
30. Sandage, A. *Astr. J.* **88**, 1159–1165 (1983).
31. Vandenberg, D. A. *Astrophys. J. Suppl.* **51**, 29–66 (1983).
32. Hodge, P. W. A. *Rev. Astr. Astrophys.* **19**, 357–372 (1981).
33. Clayton, D. D., Newman, M. J. & Talbot, R. J. Jr. *Astrophys. J.* **201**, 489–493 (1975).
34. Cahn, R. N. & Glashow, S. L. *Science* **213**, 607–611 (1981).
35. Paczynski, B. *Acta astr.* **20**, 47–58 (1970).

Models of rapidly rotating neutron stars

John L. Friedman*, James R. Ipser† & Leonard Parker*

* Physics Department, University of Wisconsin, Milwaukee, Wisconsin 53201, USA

† Department of Physics, University of Florida, Gainesville, Florida 32611, USA

Interest in rapidly rotating neutron stars has been reinforced by the discovery of the fast pulsar PSR1937+214 (ref. 1). Here we report results of the first numerical construction of rapidly rotating relativistic stars based on equations of state (EOS) proposed for neutron-star matter. Of nine EOS considered, none permits uniformly rotating equilibrium configurations for which the ratio T/W of rotational energy to gravitational energy exceeds 0.12. Thus whereas a rotating neutron star spun up by accretion or formed by collapse of a rapidly rotating dwarf can be unstable^{2–7} to modes with angular dependence $e^{im\phi}$ for $m = 3$ or 4, instability to a bar mode ($m = 2$) appears very unlikely. For the stiffest EOS and for stellar masses near the Chandrasekhar mass (baryon mass $M_0 \approx 1.4 M_\odot$), the upper limiting rotational frequencies imposed by the $m = 3$ or $m = 4$ instability are approximately equal to the frequency of the fast pulsar; for the softer EOS, the corresponding limiting frequencies are significantly larger. Uniform rotation increases the maximum gravitational mass of a model by an amount ranging from $\sim 14\%$ for the softest models to $\sim 30\%$ for the stiffest models.

Models of slowly rotating neutron stars were constructed by Hartle and Thorne⁸, and rapidly rotating relativistic configurations have been constructed, in varying degrees of approximation for idealized EOS, by several groups⁹⁻¹³. In order to build rapidly rotating neutron-star models for realistic EOS, we have adapted a code based on the method of Butterworth and Ipser¹², and constructed by Ipser for rapidly rotating relativistic polytropes. The iterative self-consistent-field method used is described in detail in ref. 12. Two computer programs were written independently (by Friedman and Parker and by Ipser), and the results are in good agreement. Iterations on models were continued until convergence to 0.1% or better was obtained. Checks on the accuracy are discussed below.

Nine of the EOS collected by Arnett and Bowers¹⁴ were used, ranging from the stiffer (M^{15} , L^{15} , N^{16} , O^{17}) to the intermediate (C^{18} , F^{19} , A^{20}) and to the softer (B^{21} , G^{22}). Because neutron stars are not expected to allow differential rotation, only uniformly rotating models were considered.

Of particular interest are upper limits on rotation and mass. Although lowering the central density of a configuration, rotation increases the mass that can be supported against collapse. The rotation itself is limited, however, by an instability of rotating self-gravitating fluids to nonaxisymmetric modes with angular dependence $e^{im\phi}$, for $m \geq (\pi G \rho)^{1/2} / \Omega$ (refs. 23-28). Here ϕ is the usual azimuthal angular coordinate, ρ is the density, and Ω is the rotational angular velocity. In realistic models the instability for large m is damped out by viscosity, and the $m=3$ or $m=4$ mode is likely to set the upper limit on rotation (and thus, indirectly, on mass as well).

In the newtonian regime, the $m=2$ mode becomes unstable when the ratio T/W of rotational kinetic energy to gravitational potential energy reaches a critical value $(T/W)_2 \sim 0.14$. This value is insensitive to the compressibility of the equation of state. In contrast, recent work on uniformly rotating polytropes, by Imamura *et al.*²⁹ and by Managan³⁰, shows that the critical values $(T/W)_3$ and $(T/W)_4$ that mark the newtonian instability points of the $m=3$ and $m=4$ modes decrease from $(T/W)_3 \approx 0.10$ and $(T/W)_4 \approx 0.08$, for homogeneous configurations, to $(T/W)_3 \approx 0.08$ and $(T/W)_4 \approx 0.06$, for $n=1$ polytropes. In fact, Imamura *et al.* find that the critical values decrease with increasing polytropic index n in such a way that the $m=3$ and $m=4$ instabilities always appear before uniformly rotating polytropes reach rotational shedding at the equator. Hence as relativity probably decreases the critical T/W slightly, and as the adiabatic indices of the proposed EOS correspond to the range $0.5 \leq n \leq 1$, we expect that $0.08 \leq (T/W)_3 \leq 0.1$ and $0.06 \leq (T/W)_4 \leq 0.08$ for neutron stars. In extrapolating to relativistic models, we anticipate inaccuracy in the critical T/W by an amount comparable to the difference in ω , the real part of normal-mode frequencies, for corresponding newtonian and relativistic models. Balbinski *et al.*³¹ found recently that the values of $\omega/(MR^{-3})$ agreed to within 5%, for fixed values of $GM/Rc^2 \leq 0.2$. Here M and R are the gravitational mass and the radius of a model.

Figure 1 exhibits a graph of rotational angular velocity Ω versus T/W for uniformly rotating relativistic models with baryon mass $M_0 = 1.4 M_\odot$. Here we take $T = \frac{1}{2} J \Omega$ and $W = M_p + T - M$, where J is the total angular momentum and $M_p = \int dV_p \rho$, with dV_p the proper volume element for the comoving observer and ρ the density of mass-energy that he measures¹². Models based on the stiffer EOS have larger radii and smaller central densities. Consequently, they are affected more by rotation: for a given angular velocity and mass, stiffer models are marked by substantially more oblateness and much larger values of T/W than are the softer, more centrally condensed models. The sequences terminate at the filled circles, where Ω reaches the keplerian angular velocity Ω_K and where matter on the equator rotates at the speed of a geodesic circular orbit.

We summarize some key features of the results displayed in Fig. 1 for $M_0 = 1.4 M_\odot$:

(1) All uniformly rotating sequences terminate, that is, reach the shedding limit where $\Omega = \Omega_K$, at values of $T/W < 0.12$. This

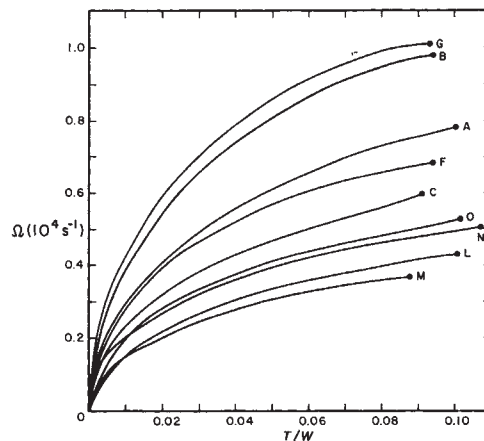


Fig. 1 Angular velocity of rotation, in units of 10^4 s^{-1} , versus the ratio of rotational kinetic energy to gravitational potential energy for uniformly rotating relativistic models with baryon mass $M_0 = 1.4 M_\odot$. Each curve is associated with a neutron-star equation of state identified in the text with the letter labelling the curve. The filled circles mark the points where equilibrium sequences terminate.

is actually independent of the mass of a model. Thus one does not expect a bar mode ($m=2$) instability to occur for uniformly rotating neutron stars.

(2) The instability points for the $m=4$ mode, at $(T/W)_4 \leq 0.08$, almost certainly occur before sequence termination. The instability points for $m=3$, at $(T/W)_3 \leq 0.1$, also probably occur in many, if not all, cases. For the stiffer EOS, the resulting upper limits on Ω are close to the frequency $\Omega_{FP} (= 0.4033 \times 10^4 \text{ s}^{-1})$ of the fast pulsar. Indeed, for EOS M, L, N, and O the values of Ω at $T/W = 0.08$ are $0.358 \times 10^4 \text{ s}^{-1}$, $0.402 \times 10^4 \text{ s}^{-1}$, $0.465 \times 10^4 \text{ s}^{-1}$, and $0.490 \times 10^4 \text{ s}^{-1}$, respectively. (Note that M is ruled out for neutron-star matter if the fast pulsar has baryon mass $M_0 \leq 1.4 M_\odot$.) For more centrally condensed models, significantly larger angular velocities (up to $\sim 1.0 \times 10^4 \text{ s}^{-1}$ for G) are required for $T/W = 0.08$. It is interesting that stiff EOS are also favoured by the redshift data for gamma-ray burst sources³².

(3) The limiting frequency depends less sensitively on gravitational mass M than on the equation of state, with $d \ln \Omega / d \ln M \sim 1$ at fixed T/W for a given equation of state. Whether the $m=3$ or $m=4$ mode (or termination of equilibrium at $\Omega = \Omega_K$) sets the upper limit on rotation also has little effect on Ω because of the flatness of the Ω versus T/W curve for $T/W > 0.08$.

(4) The observed rotation rate of the fast pulsar places a lower limit on its central mass-energy density ρ_c . For $M_0 \geq 1.4 M_\odot$, $T/W \leq 0.1$, and $\Omega = \Omega_{FP}$, we find $\rho_c \geq 3.5 \times 10^{14} \text{ g cm}^{-3}$ for the EOS studied. This refines an earlier estimate based on the bar-mode instability for homogeneous newtonian models⁵.

(5) In contrast with homogeneous configurations, none of the neutron-star models contains an ergosphere, a region within which the relativistic dragging of inertial frames causes all observers to rotate relative to the distant stars, and whose appearance is accompanied by a relativistic instability^{12,33}.

Our studies of the effect of rotation on the upper mass limits for EOS G, F, C, and L are summarized in Table 1. In each case the upper mass limit is attained for a model at the shedding limit where $\Omega = \Omega_K$, with central density $\sim 10\%$ to 20% below that of the nonrotating model with the largest mass. The greatest increase in mass is $\sim 30\%$ and corresponds to the stiffest equation of state (L). As expected, the effect of rotation decreases with increasing central concentration. Our results are consistent with those of Butterworth and Ipser¹², which show a 35% change in the mass for homogeneous relativistic models whose ratio of central pressure to density is held fixed at 0.1 as the rotation is pushed to the keplerian velocity. For the more centrally condensed models, our results are consistent with those of Hartle and Thorne⁸ for slowly rotating models, and with those of Shapiro and Lightman³⁴ for post-newtonian polytropes.

Table 1 Maximum-mass models

Equation of state	$M_{\max}/M_{\odot}$	% Increase	$M_0/M_{\odot}$	$\rho_c(10^{15} \text{ g cm}^{-3})$	$\Omega(10^4 \text{ s}^{-1})$	T/W
L	3.53	30	3.68	0.84	0.69	0.108
C	2.16	17	2.47	2.7	1.11	0.108
F	1.72	17	1.96	3.9	1.23	0.091
G	1.55	14	1.73	5.5	1.53	0.101

The first column identifies the equation of state. The remaining columns exhibit the following quantities for the rotating model with maximum gravitational mass: gravitational mass; its percentage increase over the maximum gravitational mass of nonrotating models; baryon mass; central mass-energy density; angular velocity of rotation; ratio of rotational kinetic energy to gravitational potential energy.

Table 1 also provides an estimate of the upper limiting rotation rate of a neutron star. The softest EOS permit the largest rotation rates, and for a given equation of state we find that the largest possible uniform angular velocity occurs at a termination point with Ω close to that of the maximum mass configuration. Consequently, it is evident from Table 1 that the fast pulsar is close to being able to rule out the stiffest EOS. Further, the maximum uniform Ω allowed by the softest equation of state is $\Omega_{\max} \sim 1.5 \times 10^4 \text{ s}^{-1}$, corresponding to a minimum allowed period of $\sim 0.4 \times 10^{-3} \text{ s}$. The existence of a uniformly rotating neutron star with a smaller period would rule out all proposed EOS. This result agrees with an earlier conclusion of Shapiro *et al.*³⁵ based on relativistic homogeneous configurations and Roche models.

Several checks on the accuracy of our models are available. Agreement with the Butterworth-Ipser polytrope code was obtained to six places by using a polytrope in place of tabulated EOS; in turn, the polytropic code produced $n = 3/2$ relativistic polytropes in which the values of the angular velocity of dragging agreed to within 1% with the values for Hartle's corresponding slowly rotating models³⁶. The largest angular velocity for which an equilibrium model could be constructed agreed to within 3% with the corresponding Keplerian angular velocity at the equator of the model. Finally, our results for Ω versus T/W were compared with approximate results⁴ obtained previously by using rotating newtonian polytropes to estimate the changes rotation induces in the energies and moments of inertia of spherical relativistic configurations. The comparison yielded agreement to within $\sim 15\%$, about the best that could be expected.

We thank W. D. Arnett, S. Detweiler, and L. Lindblom for helpful discussions. One of us (JRI) thanks R. Coldwell for programming assistance. The work reported here was supported in part by Grants PHY-8441263 and PHT-8300190 from the NSF.

Received 29 June; accepted 10 September 1984.

- Backer, D. C., Kulkarni, S. R., Heiles, C., Davis, M. M. & Goss, W. M. *Nature* **300**, 615-618 (1982).
- Papaloizou, J. & Pringle, J. E. *Mon. Not. R. astr. Soc.* **184**, 501-508 (1978).
- Wagoner, R. V. *Astrophys. J.* **278**, 345-348 (1984).
- Friedman, J. L. *Phys. Rev. Lett.* **51**, 11-14 (1983).
- Harding, A. K. *Nature* **303**, 683-684 (1983).
- Ray, A. & Chitre, S. M. *Nature* **303**, 409-410 (1983).
- Cowsik, R., Ghosh, P. & Melvin, M. A. *Nature* **303**, 308-310 (1983).
- Hartle, J. B. & Thorne, K. S. *Astrophys. J.* **153**, 807-834 (1968).
- Bardeen, J. M. & Wagoner, R. V. *Astrophys. J.* **167**, 359-423 (1971).
- Wilson, J. R. *Astrophys. J.* **176**, 195-204 (1972).
- Bonazzola, S. & Schneider, J. *Astrophys. J.* **191**, 273-285 (1974).
- Butterworth, E. M. & Ipser, J. R. *Astrophys. J.* **204**, 200-223 (1976).
- Butterworth, E. M. *Astrophys. J.* **204**, 561-572 (1976).
- Arnett, W. D. & Bowers, R. L. *Astrophys. J. Suppl.* **33**, 415-436 (1977).
- Pandharipande, V. & Smith, R. A. *Nucl. Phys. A* **237**, 507-532 (1975).
- Walecka, J. D. *Ann. Phys.* **83**, 491-529 (1974).
- Bowers, R. L., Gleeson, A. M. & Pedigo, R. D. *Phys. Rev. D* **12**, 3043-3055 (1975).
- Bethe, H. A. & Johnson, M. *Nucl. Phys. A* **230**, 1-58 (1974).
- Arponen, J. *Nucl. Phys. A* **191**, 257-282 (1972).
- Pandharipande, V. *Nucl. Phys. A* **174**, 641-656 (1971).
- Pandharipande, V. *Nucl. Phys. A* **178**, 123-144 (1971).
- Canuto, V. & Chitre, S. M. *Phys. Rev. D* **9**, 1587-1613 (1974).
- Chandrasekhar, S. *Phys. Rev. Lett.* **24**, 611-615 (1970).
- Friedman, J. L. & Schutz, B. F. *Astrophys. J.* **222**, 281-296 (1978).
- Friedman, J. L. *Commun. math. Phys.* **62**, 247-278 (1978).
- Lindblom, L. & Detweiler, S. L. *Astrophys. J.* **211**, 565-567 (1977).
- Comins, N. *Mon. Not. R. astr. Soc.* **189**, 233-253 (1979).
- Hiscock, W. A. & Lindblom, L. *Ann. Phys.* **151**, 466-496 (1983).
- Imamura, J., Friedman, J. L. & Durisen, R. Preprint, Los Alamos National Laboratory (1984).

- Managan, R. Preprint, Univ. Florida (1984).
- Balbinski, E., Detweiler, S., Lindblom, L., & Schutz, B. Preprint, Univ. Florida (1984).
- Lindblom, L. *Astrophys. J.* **278**, 354-368 (1984).
- Friedman, J. L. *Commun. math. Phys.* **63**, 243-255 (1978).
- Shapiro, S. L. & Lightman, A. P. *Astrophys. J.* **207**, 263-278 (1976).
- Shapiro, S. L., Teukolsky, S. & Wasserman, I. *Astrophys. J.* **272**, 702-707 (1983).
- Hartle, J. B. & Friedman, J. L. *Astrophys. J.* **196**, 653-660 (1975).

Operational approach to phase-space measurements in quantum mechanics

R. F. O'Connell

Department of Physics and Astronomy, Louisiana State University, Baton Rouge, Louisiana 70803, USA

D. F. Walls*

Institute for Theoretical Physics, University of California, Santa Barbara, California 93106, USA

Wódkiewicz¹ has derived an operational formula for a positive phase-space distribution function in quantum mechanics (see also ref. 2). Here we point out that the proposed formula is actually a special case of a two-particle Wigner distribution function in which correlations have been neglected. We present a new operational formula which includes correlations. Also, we incorporate a brief description of the role and significance of quantum distribution functions in general.

Realizing that one of the most powerful approaches to classical mechanics is through the concept of phase-space, Wigner³ introduced a quantum distribution function of positional and momentum coordinates, which provides a framework for an exact reformulation of non-relativistic quantum mechanics in terms of classical concepts. A generalization to the case of spin one-half particles has also been presented⁴.

In particular, the use of Wigner's distribution function permits one to replace a quantum-mechanical ensemble average by a classical phase-space integration. Also, as distinct from Schrödinger-Heisenberg quantum mechanics, a development of various results in powers of $\hbar$ is relatively straightforward, the limit $\hbar \rightarrow 0$ leading immediately to classical mechanics. Thus, it is not only a useful calculational tool but it also provides insights into the nature of quantum mechanics and its role in such areas as measurement theory. Similar remarks apply to other quantum-distribution functions which are in common use in many areas of non-equilibrium statistical mechanics, most particularly in quantum optics⁵. (These functions and their inter-relationships are reviewed in ref. 6.)

Wódkiewicz¹ defines his function thus:

$$P(q, p) = \iint dq' dp' W_{\psi}(q+q', p+p') W_{\phi}(q', p') \quad (1)$$

where q, q' and p, p' refer to position and momentum coordinates, respectively, and W denotes the Wigner quantum distribution function³, W_{ψ} describing the state of a laser pulse with which a detected particle interacts and W_{ϕ} describing the state of the particle after the interaction. Note that $P(q, p)$ is always positive (in contrast to $W(q, p)$ which can sometimes be negative), and that the so-called "... deficiencies of Wigner's calculation ... have been explained"².

Our main purpose here is to point out that $P(q, p)$ is not a new phase-space distribution function. It is, in fact, a special case of a two-particle Wigner distribution function in which correlations have been neglected. (Here 'particle' should be interpreted in the general sense of incorporating the systems ψ and ϕ referred to above.) The W functions appearing in equation (1) are single-particle Wigner functions and these are the kind most often discussed. However, Wigner's original paper actually introduced the general n -particle function, $W^{(n)}$ say,

* Permanent address: Department of Physics, University of Waikato, Hamilton, New Zealand.

whose properties are discussed in ref. 7. In particular, these functions obey a quantum BBGKY (Bogoliubov-Born-Green-Kirkwood-Yvon) hierarchy of equations⁸, the rate of change of $W^{(n)}$ depending not only on $W^{(n)}$ but also on the higher-order function $W^{(n+1)}$.

The two-particle Wigner function for a pure state, $W^{(2)}$ say, may be written in the form (equation (5) of ref. 3, with $n=2$):

$$W^{(2)}(q_1, q_2; p_1, p_2) = (\pi\hbar)^{-2} \iint dy_1 dy_2 \Phi^*(q_1 + y_1, q_2 + y_2) \times \Phi(q_1 - y_1, q_2 - y_2) \exp \left[\frac{2i}{\hbar} (p_1 y_1 + p_2 y_2) \right] \quad (2)$$

where Φ is the wave function for the system. If we now assume that there are no correlations between particles 1 and 2, we may write

$$\Phi(1, 2) = \psi(1)\phi(2) \quad (3)$$

from which it follows that

$$W^{(2)}(q_1, q_2; p_1, p_2) = W_\psi(q_1, p_1) W_\phi(q_2, p_2) \quad (4)$$

that is the two-particle function may be written as a product of single-particle functions. Hence, from equations (1) and (4), we deduce that

$$P(q, p) = \iint dq' dp' W^{(2)}(q + q', q'; p + p', p') \quad (5)$$

In other words, the operational phase-space distribution function $P(q, p)$ is obtained by starting with the two-particle Wigner function $W^{(2)}(q_1, q_2; p_1, p_2)$ and then (1) assuming that there is no correlation between particles 1 and 2, (2) that $q_1 = q + q' = q + q_2$, and $p_1 = p + p' = p + p_2$, and (3) integrating over all q' and p' .

We conclude that $P(q, p)$, as defined in equation (1), is a special case of a two-particle Wigner distribution function. In addition, we may regard $P(q, p)$, as defined in equation (5), as a new operational formula for a distribution function, which includes correlations.

We are not decrying the work of Wódkiewicz¹—which has the merit of deriving $P(p, q)$ from a dynamical process involving the detected particle, the detector, and the filtering device—but we are rather putting it in its proper context. In essence, it does not highlight '... deficiencies of Wigner's calculation...' but rather it displays the wide scope and power of the results presented by Wigner³.

In the same context, we point out that Wódkiewicz's function $P(q, p)$ is the same as the function $J(p, q)$ introduced by R.F.O'C. and Rajagopal⁹, in their demonstration of the fact that the '... new interpretation of the scalar product in Hilbert space...' presented by Aharonov *et al.*¹⁰ is essentially equivalent to Wigner's exact reformulation of non-relativistic quantum mechanics using distribution functions. R.F.O'C. and Rajagopal also demonstrated explicitly that $P(q, p) = J(q, p) \geq 0$ (equation (5) of ref. 9) but it is perhaps useful to point out that such a result is also contained in the earlier work of R.F.O'C. and Wigner¹¹ (see their equations (2) and (5)), who used such a result to prove that the Husimi distribution¹² (which is obtained by smoothing a Wigner distribution function with the Wigner distribution function for the ground-state of a harmonic oscillator) is always positive. The Aharonov *et al.* interpretation can also be regarded as the one underlying the stochastic phase-space formulation of non-relativistic quantum mechanics, as demonstrated by Prugovečki¹³.

All previous discussions treated pure states, but a generalization to mixed states is readily obtained. In fact, using the general result (equation (2.11) of ref. 7) that

$$\iint dq dp A(q, p) B(q, p) = (2\pi\hbar) \text{Tr}(\hat{A}\hat{B}) \quad (6)$$

where Tr denotes the trace and $\hat{A}$ and $\hat{B}$ are operators in Hilbert space and where $A(p, q)$ and $B(p, q)$ are the corresponding

Wigner phase-space functions, and also using the fact that the phase-space function corresponding to the density matrix $\hat{\rho}$ is $(2\pi\hbar)$ times the distribution function⁴, it follows that (taking $\hat{A} = \hat{\rho}_\psi$ and $\hat{B} = \hat{\rho}_\phi$)

$$\iint dq dp W_\psi(q, p) W_\phi(q, p) = (2\pi\hbar)^{-1} \text{Tr}(\hat{\rho}_\psi \hat{\rho}_\phi) \geq 0 \quad (7)$$

a result also reported by Iagolnitzer¹⁴ in work on the S-matrix description of particle states and of their measurements. In other words, the operational phase-space distribution function in the absence of correlations is always non-negative, even for mixed states.

The research of R.F.O'C. was partially supported by the US Department of Energy, Division of Materials Science, under contract DE-AS05-79ER10459 and grant DE-FG05-84ER45135 and that of D.F.W. was partially supported by the NSF under grant PHY77-27084 and the Claude McCarthy Foundation.

Received 3 July; accepted 13 September 1984.

1. Wódkiewicz, K. *Phys. Rev. Lett.* **52**, 1064-1067 (1984).
2. Maddox, J. *Nature* **308**, 601 (1984).
3. Wigner, E. P. *Phys. Rev.* **40**, 749-759 (1932).
4. O'Connell, R. F. & Wigner, E. P. *Phys. Rev. A* **30**, 2613 (1984).
5. See, for example, Drummond, P. D., Gardiner, C. W. & Walls, D. F. *Phys. Rev. A* **24**, 914-926 (1981).
6. O'Connell, R. F. in *Non-Equilibrium Quantum Statistical Physics* (eds Moore, G. & Scully, M. O.) (Plenum, New York, in the press).
7. Hillery, M., O'Connell, R. F., Scully, M. O. & Wigner, E. P. *Phys. Rep.* **106**, 121-167 (1984).
8. Balescu, R. *Equilibrium and Non-Equilibrium Statistical Mechanics*, 81, 98 (Wiley-Interscience, New York, 1975).
9. O'Connell, R. F. & Rajagopal, A. K. *Phys. Rev. Lett.* **48**, 525-526 (1982).
10. Aharonov, Y., Albert, D. Z. & Au, C. K. *Phys. Rev. Lett.* **47**, 1029-1031 (1981).
11. O'Connell, R. F. & Wigner, E. P. *Phys. Lett.* **85A**, 121-126 (1981).
12. Husimi, K. *Proc. Phys. Math. Soc. Jap.* **22**, 264-268 (1940).
13. Prugovečki, E. *Phys. Rev. Lett.* **49**, 1065-1068 (1982).
14. Iagolnitzer, D. *J. math. Phys.* **10**, 1241-1264 (1969).

The North Equatorial Countercurrent and heat storage in the western Pacific Ocean during 1982-83

Gary Meyers

CSIRO Division of Oceanography, GPO Box 1538, Hobart, Australia

Jean-Rene Donguy

ORSTOM Groupe SURTROPAC, BP A5, Noumea, New Caledonia

A sequence of extreme climate anomalies known as an El Niño Southern Oscillation episode was observed in the tropical Pacific Ocean during 1982-83^{1,2}. Models suggest that oceanic circulation has an important role in the formation of such anomalies^{1,3,4} by altering the pattern of sea-surface temperature through advection. The baroclinic structure of the upper 500 m of the ocean has been monitored routinely in the western Pacific since 1979, providing indirect current measurements by the geostrophic method. These observations, reported here, show large changes in the near-equatorial currents during 1982-83 which are consistent with currents observed in the central and eastern Pacific^{5,6}. In particular, the North Equatorial Countercurrent (NECC) flowed with 25-50% increased strength during the early phase, then weakened almost to zero flow. The West Pacific heat pool cooled by more than 1 °C. Observed changes in circulation were large enough to alter surface heat storage by advection, although other potentially important processes may not be negligible.

Baroclinic currents in the tropical ocean are largely determined by the topography of the thermocline (that is, the temperature field), in the same way that wind is determined by a map of pressure. Temperature structure has been monitored using expendable bathythermographs (XBTs) launched from volunteer observing (merchant) ships in a programme operated jointly by France and the United States since 1979⁷. The programme involves soundings to 450 m depth at intervals of ~60

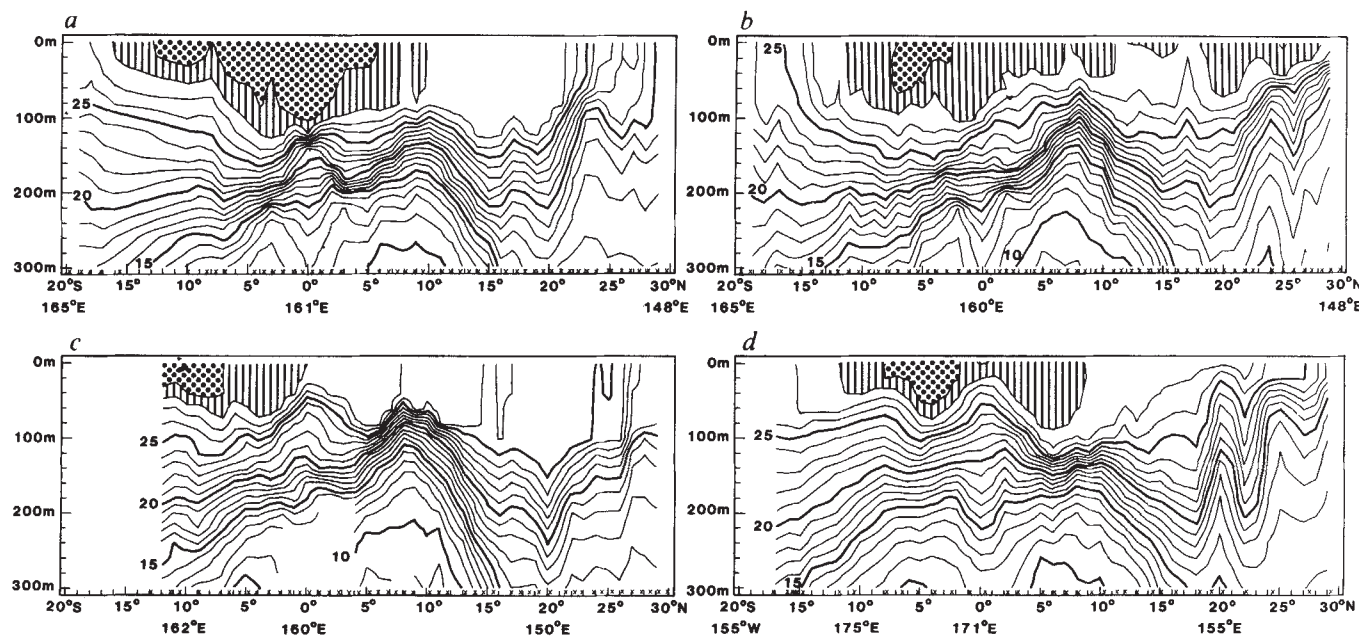


Fig. 1 Meridional temperature sections along the shipping route between New Caledonia and Japan. *a*, 2–12 January 1982; *b*, 28 July–7 August 1982; *c*, 8–17 January 1983; *d*, 1–21 June 1983. Waters with temperatures exceeding 28 °C and 29 °C are highlighted by stripes and dots, respectively. Station locations are indicated by x at the bottom of each panel.

nautical miles along the shipping route between New Caledonia and Japan, the time between cruises being typically 2–4 weeks. After careful quality control, the observations are plotted as a sequence of meridional temperature sections. Four sections chosen to show the essential features of time variability during the last El Niño episode are shown in Fig. 1.

The tropical thermocline has a series of ridges and troughs, the most persistent of which appear (Fig. 1) near 20° N, 20° S and 9° N. In the following discussion, deep levels of the thermocline are called ridges, and shallow levels troughs, in order to maintain their correspondence to ridges and troughs in sea level elevation. The major zonal equatorial currents flow parallel to the ridges and troughs over great distances across the Pacific. The ridge/trough structure varies with monthly and longer-period changes in the field of the trade winds⁸. The NECC near 3–9° N and the South Equatorial Current (SEC) near 2–20° S showed large changes during 1982–83.

The trough near 9° N on the northern side of the NECC began at a nearly normal level⁷ during January 1982 (Fig. 1*a*); it intensified by August 1982 (Fig. 1*b*), persisted until January 1983 (Fig. 1*c*), then relaxed almost completely by June 1983 (Fig. 1*d*). The slope of the 20 °C isotherm between the trough near 9° N and the ridge near 3° N is a frequently used index of NECC strength at the surface^{9,10}. A time series of monthly values (Fig. 2*a*) was prepared by averaging the depth of the 20 °C isotherm within $\pm 1.5^\circ$ latitude of the ridge or trough, then taking the difference (ΔD_{20}). We used this isotherm for the index because it is near the depth of maximum vertical temperature gradient in the NECC.

The index of relative geostrophic current strength derived in this way is a rather smooth variable, dominated by low-frequency seasonal and interannual fluctuations. The eastward surface flow of the NECC was stronger than normal during June to December 1982; at approximately the same time, enhanced eastward flow was observed in the central⁵ and eastern⁶ Pacific. Thereafter, the NECC diminished to an extremely weak level by March 1983 and remained weak until the end of our record (June 1983). The signal during 1982–83 suggests an annual oscillation; however, earlier studies^{11,12} have shown that seasonal oscillations are very irregular in amplitude and phase in this region. In association with the changing strength of the current, the ridge on the southern side varies in latitude between the Equator (Fig. 1*b*) and 5° N (Fig. 1*d*) during the episode. Variability in

the strength of the NECC during earlier episodes has been inferred from sea level¹¹; however, continuous spatial sampling by XBTs is needed to locate the latitude of ridges and troughs and to determine the strength of the full current. The ΔD_{20} index (Fig. 2*a*) shows that the surface current was 25–50% stronger than normal during the early part of the El Niño episode.

Change in the SEC was also observed on the temperature sections. The SEC was a very stable feature from 1979 until mid-1982⁷. Its strongest surface flow usually appeared between the Equator and 10° S, as shown by the downward slope of all isotherms in this latitude band (Fig. 1*a*). The current shifted southward to 10–20° S (Fig. 1*d*) during the 1982–83 episode, as reported in more detail elsewhere⁷.

Another aspect of the temperature field in the western Pacific is a pool of very warm water near the Equator. Its normal structure is seen in January 1982 (Fig. 1*a*) as a pool of 28–29 °C water reaching more than 100 m deep in places. Typically, the near-equatorial (5° N to 5° S) portion of the 29 °C pool extends eastward a short distance beyond the New Caledonia–Japan section (160° E) and westward to Papua New Guinea¹³ (about 140° E). The heat pool shifted into the central Pacific during 1982–83 and during an earlier El Niño episode^{14,15}, while surface temperatures in the west decreased. The recently recognized potential importance of western Pacific temperature in controlling the Southern Oscillation^{1,16} suggests that it is essential to document its variability and to understand the processes that control it.

The sequence of temperature sections (Fig. 1) shows changes in the heat pool very clearly. After the normal conditions observed in January 1982 (Fig. 1*a*), the 29 °C pool was depleted in the near-equatorial band by August 1982 (Fig. 1*b*). The cooler 28 °C pool was almost depleted by January 1983 (Fig. 1*c*). Near the end of the episode, in June 1983 (Fig. 1*d*), the heat pool is being re-established in the near-equatorial band, but its volume is still substantially below normal. The heat loss is not restricted to the near-equatorial band, but in fact appears in the complete sequence of sections (not presented) as a prominent interannual signal throughout the latitude band 20° N–15° S, superimposed on the usual seasonal cycle. A monthly index of near-equatorial heat content was derived from the complete sequence of temperature sections by averaging the temperature between the sea surface and the depth of the 26 °C isotherm in the latitude band 1° S–5° N (Fig. 2*b*). Documenting heat storage in this way

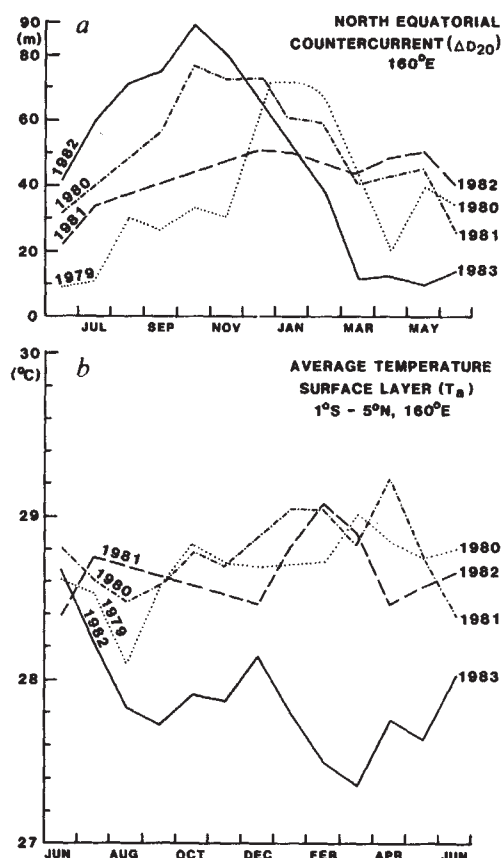


Fig. 2 *a*, Index of surface current speed in the North Equatorial Countercurrent estimated by the difference in depth of the 20 °C isotherm from the southern to the northern side of the current. *b*, Index of temperature in the mixed layer estimated by the vertically averaged temperature between the surface and the depth of the 26 °C isotherm. In *a* and *b* the values for August to November 1981 were interpolated.

reduces noise associated with adiabatic vertical motion due to variability of currents¹⁷. The index (T_a) shows that heat storage during the second half of 1982 and early 1983 was very abnormal compared with the previous 3 yr in that the western Pacific heat pool had cooled by 1.3 °C from June 1982 to March 1983. The cooling occurred relatively rapidly during two stages, June–August 1982 and January–March 1983. Note that the cooling stages correspond to the winter months of either hemisphere. The average cooling over the 9-month period implies a local heat flux of -20 W m^{-2} , taking into account that the average depth of 26 °C was 90 m during that time.

Theoretically, the processes which control heat storage in the surface layer are heat fluxes through the sea surface, entrainment of cold water from the thermocline and advection. Methods of assessing the magnitude of these processes from existing data have errors of the order of 20 W m^{-2} , which prevents the determination of a heat budget to account for the observed cooling. Nevertheless, the plausibility of the three processes can be discussed. Surface fluxes and the possibility of colder water from deeper layers replacing displaced warm surface waters through upwelling and mixing will be discussed in a subsequent letter. Advection is discussed here.

The observed increase in eastward surface current (Fig. 2*a*) during the early part of the 1982–83 episode is qualitatively consistent with a model of warm pool migration into the central Pacific¹. The model attributes increasing temperature in the central Pacific to anomalous eastward advection. By assuming an advective model, we can estimate the rate of this process from the XBT observations. A volume transport of $10 \times 10^6 \text{ m}^3 \text{ s}^{-1}$ is required to remove the 110-m thick (Fig. 1*a*) layer of

28 and 29 °C water from the region 1° S–5° N, 140°–160° E over 6 months (June–December 1982). This volume transport could be largely accommodated by the 25–50% increase in strength of the NECC (Fig. 2*a*), and possibly other eastward currents at the Equator. In this sense, our observations are consistent with the model of temperature increase in the central Pacific.

But can the western Pacific cool down by advection? Mass continuity is partly maintained by the 50-m rise in the thermocline, which leaves approximately half of the volume transport ($5 \times 10^6 \text{ m}^3 \text{ s}^{-1}$) to accomplish the -20 W m^{-2} cooling (Fig. 2*b*) by horizontal advection. This cooling rate can be achieved by invoking a temperature gradient of 1.1 °C over a distance of 2,000 km along the trajectory of the current (assuming the width of the current is 5° of latitude). However, what is the source of the cooler water? Waters entering the heat pool in the North and South Equatorial Currents are a few degrees cooler during the winter months of either hemisphere than water leaving in the NECC. The cooler waters of the NEC and SEC flow to the NECC by way of a complex system of meridional currents near the western boundary, so that the temperature gradient of 1.1 °C per 2,000 km often exists along the trajectory. In support of a purely advective model, climatological maps of surface temperature¹³ for winter of either hemisphere show an abundance of relatively cool waters near the head waters of the meridional currents. Furthermore, documentation of basinwide surface temperature anomalies during 1982–83¹⁸ using a comprehensive XBT data set shows a dominant spatial pattern of western Pacific cooling extended eastward along the NECC axis. Thus, the purely advective model is possible for the western Pacific, as it is for the central Pacific¹, despite small horizontal temperature gradients in the region.

This research was partially supported by NSF grants (to G.M.) as part of the PEQUOD program. We gratefully acknowledge this support. J. Church and C. Fandry read an early version of this paper and suggested important improvements.

Received 23 July; accepted 5 September 1984.

- Gill, A. E. & E. M. Rasmusson *Nature* **306**, 229–234 (1983).
- Philander, S. G. H. *Nature* **305**, 16 (1983).
- Philander, S. G. H. *Nature* **302**, 295–301 (1983).
- Cane, M. A. *Science* **222**, 1189–1195 (1983).
- Firing, E., Lukas, R., Sadler, J. & Wyrtki, K. *Science* **222**, 1121–1123 (1983).
- Halpern, D., Hayes, S. P., Leetmaa, A., Hansen, D. V. & Philander, S. G. H. *Science* **221**, 1173–1175 (1983).
- Meyers, G. & Donguy, J. R. *Trop. Ocean-Atmos. Newslett.* **2**, 6–7 (1980); **16**, 8–9 (1983); **21**, 8–9 (1983); **27**, 10–11 (1984).
- Busalacchi, A. J., Takeuchi, K. & O'Brien, J. J. *J. geophys. Res.* **88**, 7551–7562 (1983).
- Wyrtki, K. & Kendall, R. *J. geophys. Res.* **72**, 2073 (1967).
- Wyrtki, K. *et al. Science* **211**, 22–28 (1981).
- Wyrtki, K. *J. phys. Oceanogr.* **9**, 1223–1231 (1979).
- Meyers, G. *J. phys. Oceanogr.* **12**, 1161–1168 (1982).
- Reynolds, R. W. *A Monthly Averaged Climatology of Sea Surface Temperature* (US National Oceanic and Atmospheric Administration Tech. Rep. NWS 31, 1982).
- Meyers, G., Donguy, J. R. & Cutchin, D. in *1982–83 El Niño Southern Oscillation Workshop*, Miami (ed. Witte, J.) 33–52 (Nova University Press, Dania, Florida, 1983).
- Donguy, J. R., Dessier, A., Eldin, G., Morliere, A. & Meyers, G. *J. mar. Res.* **42**, 103–121 (1984).
- Nicholls, N. *Mon. Weath. Rev.* **112**, 424–432 (1984).
- Stevenson, J. A. & Niiler, P. P. *J. phys. Oceanogr.* **13**, 1894–1907 (1983).
- White, W. B., Meyers, G., Donguy, J. R. & Pazan, S. *J. phys. Oceanogr.* (submitted).
- Niiler, P. P. & Stevenson, J. S. *J. mar. Res.* **40**, 465–485 (1982).

Iridium in Mississippi River suspended matter and Gulf of Mexico sediment

F. D. Fenner & B. J. Presley

Oceanography Department, Texas A&M University, College Station, Texas 77843, USA

The suggestion¹ that a meteorite impact at the end of Cretaceous time (the Cretaceous–Tertiary boundary) caused the extinction of many groups of organisms and left an anomalous iridium concentration in sediments has stimulated interest in both the extinctions and in the geochemistry of Ir (refs 2–6). The geochemical behaviour of Ir is not well understood, partly because of the scarcity of concentration data for common rocks, soils and waters. Here

we report Ir abundances in Mississippi River suspended matter (MRSM) and Gulf of Mexico (GOM) sediments and suggest that these values may be representative of much of the upper continental crust (that accessible to rock weathering). These data imply that the processes leading to the well known enrichment of iridium in deep-sea sediments⁷ are similar to those which enrich elements such as Ni and Co.

Most of the samples used in this study were collected by Trefry⁸ who determined their major and minor inorganic constituents. Bottom sediment was collected on the shelf, slope and abyssal plain of the Gulf of Mexico by a conventional 7-cm diameter gravity core equipped with a polybutyrate liner. Suspended matter was separated from water collected by hand-dipping with a plastic bucket or with a plastic-lined pump, either by filtration through 0.4 µm pore diameter membrane filters or by simple gravity settling and centrifuging. No difference in chemistry could be attributed to the different collection techniques. Samples were taken at Head of Passes, ~20 km upstream from the mouth of the Mississippi River, from locations further upriver and from the freshwater-seawater mixing zone at the river mouth⁹.

Neutron activation analysis is highly sensitive to Ir, but due to interferences from other elements, this technique must be combined with radiochemical separation for samples such as ordinary soils and sediment^{3,10}.

In the present work, freeze-dried, finely ground samples (0.5 g) were accurately weighed into acid-cleaned quartz vials which were then heat-sealed. These, together with flux monitors and standards, were exposed to a flux of 1×10^{13} neutrons cm⁻² s⁻¹ for 36-h periods at the Texas A&M University reactor. After irradiation, the samples were allowed to cool for 15 days and the Ir was then separated essentially as described by Gijbels¹⁰. The method involves dissolution of the sample by fusion with sodium hydroxide and sodium peroxide, separation of Ir on an anion-exchange resin, elution, precipitation of IrO₂ and finally filtration and counting. After counting the ¹⁹²Ir induced in the

samples, the filters were re-irradiated for 1 min to activate Ir added as a carrier at the dissolution step. A second count was then used to determine chemical yield, which was generally 75–80%. Reagent blanks and granite rocks were below the detection limit of 0.01 parts per 10⁹ (p.p.b.) whereas a C-T boundary clay sample from Stevens Klint and a USGS standard rock (DTS-1) gave 45 and 61 p.p.b. respectively, in agreement with published results^{3,11}.

Metal concentration data for MRSM and the Mississippi submarine sediment are given in Table 1. When the other metal concentrations were plotted against Fe, high linear correlation coefficients were found (Table 3), as had been noted in previous work on samples from this geographical area¹². Iridium, which had not been determined previously on Mississippi River samples, and apparently not on any river samples, followed the same concentration pattern as the other elements.

The Fe-metal regression equations calculated from MRSM data obtained in the present study were used with Trefry's⁸ average Fe concentration to calculate concentrations of other key elements. The calculated concentrations and those actually measured by Trefry⁸ are given in Table 2. The measured values generally fall within the 95% confidence interval of the calculated values, except for Co where analytical problems are suspected. It appears, then, that some trace metal concentrations can be reliably predicted from Fe concentration for MRSM and Gulf of Mexico sediments, and that Ir is one of the predictable elements.

Success in using Fe data to predict trace metal concentrations in the small MRSM data set led us to generalize the calculations. An average value for Fe in the continental crust¹³ was used with the MRSM regression equations to predict crustal abundances of various trace elements. The results of these calculations (Table 3) show that the predicted values for Cr, Ni, Sc and La agree with published estimated values. Literature Co and V values fall slightly outside the 95% confidence limits of the regression prediction. The predicted Th value is 72% higher than the value

Table 1 Elemental concentrations in Mississippi River suspended matter and delta sediment

Sample	Al (%)	Fe (%)	V (p.p.m.)	Mn (p.p.m.)	Sb (p.p.m.)	Co (p.p.m.)	Zn (p.p.m.)	Cu (p.p.m.)	Ni (p.p.m.)	Cr (p.p.m.)	La (p.p.m.)	Sc (p.p.m.)	Th (p.p.m.)	Ir (p.p.b.)
Delta sediment														
Station 8 (20–30 cm)	7.50	4.23	147	739	1.4	12	134	47	46	114	44	15.9	12.6	0.050
Station 34 (0–10 cm)	7.93	4.53	164	640	1.4	13	130	51	40	107	41	16.7	13.3	0.061
Station 32 (20–30 cm)	7.25	3.65	194	1,480	1.4	17	107	40	56	86	32	14.6	10.5	0.040
Station 36 (20–30 cm)	5.19	2.32	81	344	1.6	9	70	15	14	67	34	8.8	9.6	0.035
Station 28 (20–30 cm)	6.42	3.36	109	732	2.0	10	110	55	36	99	39	12.9	11.3	0.045
Suspended matter														
Head of Passes	9.77	5.97	190	950	1.7	15	196	72	69	110	41	20.7	14.8	0.085
St Francisville, Louisiana	7.94	4.71	147	1,370	2.8	15	171	63	51	95	39	16.6	13.0	0.088
Arkansas 82	4.39	1.49	40	231	4.6	7	60	12	13	42	23	5.5	7.2	0.019
Mixing zone (1%)	6.55	3.64	109	714	4.5	12	132	65	24	78	27	13.2	10.5	0.047
Mixing zone (4%)	7.33	4.00	139	937	2.4	13	164	70	32	88	30	14.6	13.0	0.043
Mixing zone (24%)	5.60	3.62	102	1,180	—	14	133	56	24	—	27	12.4	9.9	0.054

Sampling locations are given in ref. 9.

Table 2 Average concentrations for selected trace elements in MRSM, concentrations calculated from Fe and Fe-metal regression equations and per cent difference

Element	Average MRSM concentration* (p.p.m.)	Predicted concentration (p.p.m.)	Δ (%)
Ni	55	45.5 ± 17†	–17
Co	21	14.3 ± 3	–32
Cr	79	99 ± 29	+20
Zn	193	156 ± 31	–19
V	150	151 ± 19	–1

* From ref. 8.

† 95% confidence interval.

Table 3 CCA, CCA predicted from MRSM regression equation and Fe CCA, per cent difference and correlation coefficient (r)

Element	CCA* (p.p.m.)	Predicted (p.p.m.)	Δ (%)	r
Cr	100	113 ± 31†	+13	0.90
Ni	75	58 ± 18	–23	0.92
Co	22	15.5 ± 3	–30	0.90
V	135	185 ± 29	+37	0.97
Th	8.5	14.6 ± 2.0	+72	0.94
Sc	20	20.2 ± 1.6	+1	0.99
La	35	41.4 ± 1.5	+18	0.66
Ir‡	0.025§	0.081 ± 0.025	+225	0.91

* From ref. 13.

† 95% confidence interval.

‡ Value in p.p.b.

§ From ref. 7.

Table 4 Elemental concentrations in Gulf of Mexico sediments (CaCO₃-free)

Core notation	Fe (%)	V (p.p.m.)	Mn (p.p.m.)	Sb (p.p.m.)	Co (p.p.m.)	Zn (p.p.m.)	Cu (p.p.m.)	Ni (p.p.m.)	Cr (p.p.m.)	Ag (p.p.b.)	Ir (p.p.b.)	CaCO ₃ (%)
OB-1	4.28	158	1,390	1.6	18.2	118	34	46	105	200	0.11	17.7
OB-2	—	169	1,290	—	—	123	38	43	—	320	0.088	20.2
OB-4	—	154	1,650	—	—	129	37	48	—	220	0.063	17.6
OB-5	4.52	168	2,260	3.7	19.0	124	37	56	105	220	0.074	19.2
OB-6	4.42	164	1,390	1.4	18.0	122	37	47	114	156	0.072	19.0
OB-7	4.31	123	1,160	2.6	11.6	69	24	48	75	194	0.066	12.2
OB-8	4.01	164	1,280	1.9	15.4	115	38	46	96	220	0.079	20.1
OB-9	4.05	176	1,000	2.0	12.8	—	44	53	92	—	0.085	20.1
OB-10	4.53	178	1,450	1.6	16.1	—	45	66	104	—	0.067	21.6
GOM 4	4.50	163	774	1.7	16.4	—	29	56	111	—	0.061	4.5
GOM 5	3.50	—	—	1.4	21.9	—	47	60	106	—	0.053	—
GOM 6	2.92	—	3,320	2.9	29.2	—	—	87	83	—	0.097	25.0

OB 1-10, Individual cores from Orca Basin²².
 GOM 4, Gulf of Mexico continental rise core⁸.
 GOM 5, Gulf of Mexico abyssal plain core⁸.
 GOM 6, Gulf of Mexico Sigsbee Knoll core⁸.

in the literature and the predicted Ir concentration 225% higher. We suggest that the large Ir discrepancy is due to an underestimation of Ir continental crustal abundance (CCA) by previous workers.

One widely used estimate for Ir CCA, based on a study of Canadian Precambrian shield rocks¹⁴, gives a value of 0.02 p.p.b. Another estimate¹⁵ derives from averaging tholeiitic basalt values and those of calc-alkaline rocks; this yields 0.03 p.p.b. The weathered residue carried by the Mississippi River is enriched in Ir relative to Fe compared with these two rock types and it may be more representative of material being transported to the ocean.

The Mississippi River basin has an area of $>3 \times 10^6$ km² and includes a good cross-section of rock types. Its Ir load should be typical of large rivers. It is not contaminated by anthropogenic Ir as is shown by century-old sediment from river delta cores. During its long transport, the weathered rock residue carried by the river is thoroughly homogenized. Based on an analysis of 40 samples collected over an 18-month period (1974-75) by our research group^{8,16,17}, suspended matter from the lower Mississippi River is chemically similar to average continental rocks for Al, Fe, Mn and several trace elements, a conclusion confirmed by the data from the present study (Table 3). Furthermore, the MRSMS varies little chemically with season or year⁸, despite very large variations in river flow and suspended load.

The crustal abundance of Ir based on MRSMS, the MRSMS regression equation for Fe versus Ir, and Krauskopf's¹³ average Fe is 0.081 p.p.b.—a value more than three times the 0.025 p.p.b. estimate reported recently^{7,18}. The higher estimate of Ir CCA suggested here would lower the difference between continental rocks and average deep ocean sediment^{7,19,20} from a factor of 12 to a factor of 4. Although a four-fold enrichment must be explained, it is close to the enrichments shown by such elements as Co, Cu and Ni.

Extraterrestrial material (ETM) is greatly enriched in Ir, but it is unlikely that this material makes a large contribution to the Ir in average deep-sea sediments. Crockett and Kuo⁷ used the mass accretion rate for ETM and its Ir concentration to calculate the percentage of extraterrestrial Ir in deep-sea cores they had analysed. For a core with a sedimentation rate of 3.6 cm kyr⁻¹, only 1% of the Ir could have come from ETM, and only 17% for a core which accumulated at 0.33 cm kyr⁻¹. Sedimentation rates for the GOM cores used in our study ranged from several centimetres to several metres per 1,000 yr, making ETM negligible.

Many explanations for the enrichment of trace elements in deep-sea sediments have been offered²¹, but no consensus has been reached. Our data for riverine and near-shore sediments (Tables 1 and 4) show similar concentration trends for Ir, Co and Ni. Furthermore, deep-sea sediments are enriched over continental and near-shore materials by similar amounts for Ir,

and such elements as Co and Ni. Thus, it seems likely that similar processes work to enrich Ir and the more common trace elements. These processes are unlikely, however, to produce the extreme enrichments found at the Cretaceous-Tertiary boundary.

Received 30 April; accepted 18 September 1984.

1. Alvarez, L. W., Alvarez, W., Asaro, E. & Michael, H. V. *Science* **208**, 1095-1108 (1980).
2. Kyte, F. T., Zhou, Z. & Wasson, J. T. *Nature* **288**, 651-656 (1980).
3. Ganapathy, R. *Science* **209**, 921-923 (1980).
4. Smit, J. & Hertogen, J. *Nature* **285**, 198-200 (1980).
5. Keith, M. L. *Geochim. cosmochim. Acta* **46**, 2621-2637 (1982).
6. Officer, C. B. & Drake, C. L. *Science* **219**, 1383-1386 (1983).
7. Crockett, J. H. & Kuo, H. Y. *Geochim. cosmochim. Acta* **43**, 831-842 (1979).
8. Trefry, J. H. thesis, Texas A&M Univ. (1977).
9. Fenner, F. D. thesis, Texas A&M Univ. (1983).
10. Gijbels, R., Millard, H., Desborough, G. & Bartel, A. *Activation Analysis in Geochemistry and Cosmochemistry* (eds Bronfeldt, A. & Steinhilber, E.) 359-370 (Universitetsforlaget, Oslo, 1971).
11. Gladney, E. S., Burns, E. & Roelandts, I. *Geostand. Newsl.* **7**, 1 (1983).
12. Trefry, J. H. & Presley, B. J. *Environ. Geol.* **1**, 283-294 (1976).
13. Krauskopf, K. B. *Introduction to Geochemistry* (McGraw-Hill, New York, 1979).
14. Shaw, D. M., Dostal, J. & Keays, R. R. *Geochim. cosmochim. Acta* **40**, 73-83 (1976).
15. Greenland, L. P., Gottfried, D. & Tilling, R. I. *Chem. Geol.* **14**, 117-122 (1974).
16. Trefry, J. H. & Presley, B. J. *Marine Pollutant Transfer* (eds Windom, H. & Duce, R.) 39-76 (Lexington Books, 1976).
17. Presley, B. J., Trefry, J. H. & Shokes, R. F. *Water, Air, Soil Pollut.* **13**, 481-494 (1980).
18. Barker, J. L. & Anders, E. *Geochim. cosmochim. Acta* **32**, 627-645 (1968).
19. Harris, C. H., Crockett, J. H. & Stainton, M. *Geochim. cosmochim. Acta* **32**, 1049-1056 (1968).
20. Kyte, F. T. & Wasson, J. T. *Lunar Planet. Science Conf.* **13**, 411, Houston (1982).
21. Chester, R. & Aston, S. R. *Chemical Oceanography* (eds Riley, J. P. & Chester, R.) 281-390 (Academic, London, 1976).
22. Sheu, D. D. thesis, Texas A&M Univ. (1983).

Evaporation from forests in cloud enhances the effects of acid deposition

M. H. Unsworth

Institute of Terrestrial Ecology, Bush Estate, Penicuik, Midlothian EH26 0QB, UK

Declining growth rates and injury to foliage have been observed in forests at high elevations in areas of Europe and North America^{1,2}; in such areas the trees are regularly enveloped in cloud and mist. Concentrations of dissolved chemicals in cloud droplets are often much larger than in rain, so the capture of cloud droplets by leaves may give rise to significant rates of 'occult deposition'³ of acidic pollutants and heavy metals on forests^{3,4}. I point out here that evaporation of intercepted cloudwater from forest canopies, proceeding simultaneously with occult deposition, can lead to chemical concentrations on leaf surfaces that are substantially larger than those measured in cloud drops themselves. These large chemical concentrations may be a cause of the observed injury in forests where wind-driven cloud is common.

This analysis is concerned only with deposition on uniform and extensive tree canopies; deposition and evaporation rates

Table 1 Calculations from equation (2) of the influence of wind speed on the deposition rate of wind-driven cloud onto a pine forest

Wind speed u (m s^{-1})	2	4	6	8
Friction velocity u_* (m s^{-1})	0.59	1.19	1.78	2.38
Deposition rate D ($\text{mg m}^{-2} \text{s}^{-1}$)	17.4	35.4	52.8	70.8

Calculations are based on a cloud-water concentration 0.1 g m^{-3} and on the logarithmic wind profile observed in neutral stability⁸, that is $u = u_* k^{-1} \ln(z - d/z_0)$ where k is Von Karman's constant (0.41), d is the zero plane displacement (m), z_0 is the roughness length (m) and z is height above the ground. Values of u are referred to the top of the canopy ($z = 15.5 \text{ m}$), and empirical values derived for Thetford forest⁸ were $d = 11.8 \text{ m}$ and $z_0 = 0.93 \text{ m}$.

from isolated trees and shrubs may be much larger than those estimated here but the principles concerning surface concentrations would apply equally well. When the rate of deposition of cloud water to a unit ground area of forest is D ($\text{g m}^{-2} \text{s}^{-1}$), the rate of deposition of a substance with concentration C_d ($\mu\text{g g}^{-1}$) in cloud drops is DC_d ($\mu\text{g m}^{-2} \text{s}^{-1}$). To estimate the rate of deposition of a non-volatile substance in cloud water it is sufficient to determine D and C_d (ref. 5). But if some of the deposited water evaporates from the vegetation at the same time as occult deposition takes place, the concentration of a substance in the water films on leaf surfaces will be larger than the concentration in cloud drops. When studying the effects of acid cloud on plants it is the surface film concentration C_1 rather than the cloud drop concentration C_d which is likely to be associated with leaf responses. I use results from several micro-meteorological experiments to estimate the concentration ratio C_1/C_d and hence to illustrate how the pH of acid solution on leaf surfaces may differ substantially from that in cloud drops.

Assuming that the substance dissolved in cloud drops is non-volatile but that deposited cloud water evaporates at a rate E ($\text{g m}^{-2} \text{s}^{-1}$) the concentration ratio is given by

$$C_1/C_d = D/(D - E) \quad (1)$$

Clearly as the difference $(D - E)$ tends to zero then substantial surface concentrations can arise. To investigate whether such occasions are likely, the values of D and E when forests are enveloped in cloud need to be estimated.

There have been very few direct measurements of D , but the limited results over forests⁶ and grassland⁷ show that measured values are close to the maximum rates predicted from a knowledge of wind speed and turbulence over the canopy and the liquid water content W (g m^{-3}) of the cloud⁷, that is

$$D \cong u_*^2 W/u \quad (2)$$

where u is the wind speed at the top of the canopy and u_* is the friction velocity, a measure of the speed of turbulent mixing.

Table 1 shows values of D calculated from equation (2) using wind-speed parameters derived over Thetford forest (a mixed forest of *Pinus sylvestris* and *Pinus nigra* var. *maritima* 15.5 m high⁸) and a typical value $W = 0.1 \text{ g m}^{-3}$.

The evaporation rate E from a wet vegetation canopy depends on the energy A available for heat transfer, the humidity of the air and the efficiency of vapour transfer from the surface to the free atmosphere. Over wet forests the aerodynamically-rough surface makes E depend strongly on atmospheric humidity. In large clouds, the air is often supersaturated but in thin cloud and mist the relative humidity h can be slightly below 100%. When forests are enveloped in cloud, A is unlikely to exceed 100 W m^{-2} .

Stewart measured evaporation from Thetford forest when the canopy was wet⁹ and showed that his observations were consistent with calculations based on the Monteith-Penman equation¹⁰.

Figure 1, based on this equation shows the dependence of evaporation rate E from the wet forest on relative humidity at air temperature 10°C , and at three values of available energy. Also shown are comparable estimates of deposition rates D at

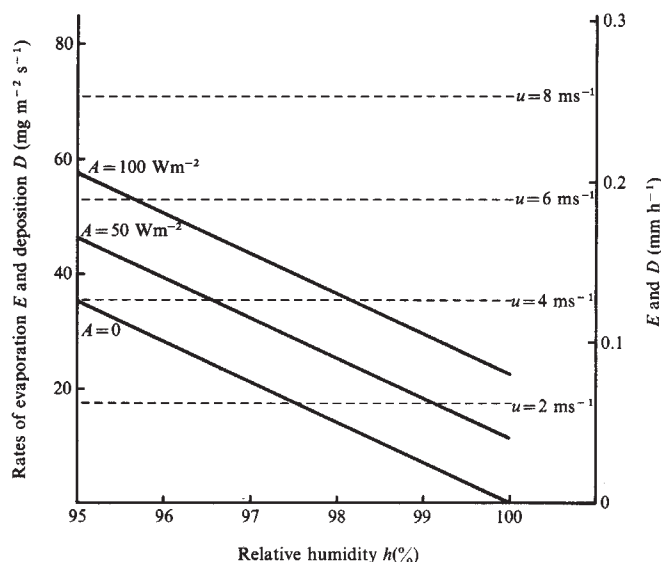


Fig. 1 The dependence of evaporation rate E (solid line) from a wet forest on relative humidity h for three values of available energy A ; and the deposition rates D (dashed line) of cloud water over the same forest, calculated from equation (2) for four wind speeds u . The values of E are calculated using the Monteith-Penman equation,

$$E = \frac{1}{\lambda} \frac{\Delta A + \rho_c p_e (1 - h)/r}{\Delta + \gamma}$$

where λ is the latent heat of vaporization of water, Δ is the rate of change of saturation vapour pressure with temperature at air temperature, ρ_c is the volumetric specific heat of dry air, p_e is the saturation vapour pressure at air temperature, h is the relative humidity, A is the available energy, γ is the psychrometric constant and r is the aerodynamic resistance to transfer of water vapour from the canopy. It was assumed that air temperature was 10°C and the aerodynamic resistance r was taken as 6 s m^{-1} , the average value derived from measurements at Thetford forest⁹.

four wind speeds. Figure 1 demonstrates that when the atmosphere is only slightly unsaturated, when there is significant available energy, and/or when wind speeds are low, values of E and D may be of a similar magnitude. In these circumstances, equation (1) predicts that for non-volatile material the concentration ratio C_1/C_d may be large.

Figure 2 illustrates the influence of atmospheric humidity on the hydrogen-ion concentration C_1 in canopy surface water for the hypothetical case of deposition on a pine forest of cloud water with an initial pH of 3.5, (typical for an episode of acid deposition¹¹). Figure 2 demonstrates that surface pHs below 3, known to visibly injure leaves and reproductive structures¹², could be generated during acid cloud deposition, even though the cloud water itself was insufficiently acid to injure vegetation.

The calculations for Fig. 2 assume that (1) acidity in the cloud water is associated with non-volatile components, and (2) no chemical reactions take place in surface film to modify C_1 . The first assumption is likely to be valid for cloud at distances of more than a few hundred km from pollution sources; in such cases the acidity is primarily associated with sulphuric and nitric acids¹³, gas concentrations of SO_2 are very low, and influences of volatile sulphite and dissolved SO_2 on acidity would be negligible¹⁴. The second assumption implies that Fig. 2 represents the maximum surface acidity directly attributable to cloud water deposition, because, in reality, there may be ion exchange between leaf tissue and surface solution which would influence the surface water acidity. The extent to which ion exchange would alter the pH from values calculated in Fig. 2 would depend on plant species, age, and previous pollution exposure^{12,15}.

Equation (1) implies that as $D - E$ tends to zero, C_1/C_d tends to infinity. In practice evaporation would cease when the solution on the leaves became so concentrated that the equilibrium

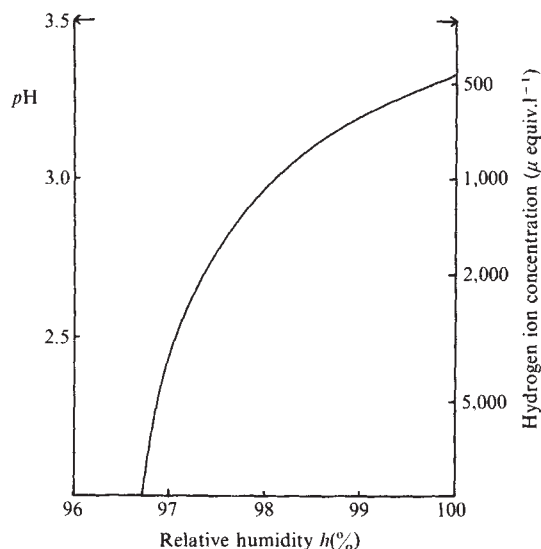


Fig. 2 The dependence of the mean concentration of hydrogen ion (pH) in surface water on a canopy on relative humidity, for typical conditions of cloud deposition and evaporation over a pine forest: $A = 50 \text{ W m}^{-2}$, $u = 4 \text{ m s}^{-1}$, air temperature 10°C (see Fig. 1). The calculated curve refers to cloud drops with initial pH 3.5; this value is indicated on the concentration axis.

vapour pressure over it was reduced by the influence of the solutes to the vapour pressure of the surrounding air. However, over the pH range shown in Fig. 2 the limiting stage is not likely to be approached. For example, the equilibrium relative humidity of air over sulphuric acid solution at a pH of 2 is $\sim 99.97\%$. Consequently, increasing solute concentration would not significantly alter the conclusions drawn here.

These calculations suggest that the effects of acid cloud or mist deposition on vegetation depend not only on chemical concentrations in the drops but also on the weather factors and vegetation characteristics that influence deposition and evaporation. Simultaneous direct measurements of deposition and evaporation have not been made at mountain sites, but Hori¹⁶ reported many cases where evaporation was substantial from forests in sea fog. To test the hypotheses developed here, it is necessary to investigate whether evaporation of drops in the air close to the canopy limits deposition or whether turbulent transfer is sufficiently effective to maintain deposition rates assumed here. This note draws attention to the mechanism that exists for increasing the concentration of liquid on leaf surfaces whenever there is simultaneous deposition and evaporation. The more obvious concentrating effect when deposition ceases and leaf surfaces dry by evaporation may also be important, but is transient whereas the mechanism described here may persist during long periods of cloud cover. In laboratory experiments to simulate effects of acid mist the ideas discussed here emphasize the need to ensure either that evaporation does not occur at rates which enhance the surface acidity, or that surface concentrations of chemicals are measured and described.

Received 18 June; accepted 21 August 1984.

1. Rehfuess, K. *Forstwissenschaftlichen Centralblatt* **102**, 167-181 (1983).
2. *The Acid Deposition Phenomenon and its Effects* vol. 2, 3-37-3-41 (EPA-600/8-83-0168, US Environmental Protection Agency, Washington DC, 1983).
3. Dollard, G. J., Unsworth, M. H. & Harvey, M. J. *Nature* **302**, 241-243 (1983).
4. Lovett, G. M., Reiners, W. A. & Olson, R. K. *Science* **218**, 1303-1304 (1984).
5. Lovett, G. M. *Atmos. Envir.* **18**, 361-371 (1984).
6. Huzioka, T., Tabata, T. & Matsumura, N. in *Studies on Fogs* (ed. Hori, T.) 205-220 (Tanne Trading Co., Sapporo, 1953).
7. Dollard, G. J. & Unsworth, M. H. *Atmos. Envir.* **17**, 775-780 (1983).
8. Oliver, H. R. Q. *J. R. met. Soc.* **97**, 548-553 (1971).
9. Stewart, J. B. *Wat. Resour. Res.* **13**, 915-921 (1977).
10. Monteith, J. L. *Principles of Environmental Physics*, 241 (Edward Arnold, London, 1973).
11. Fowler, D. & Cape, J. N. *Atmos. Envir.* (in the press).
12. Jacobson, J. S. in *Ecological impact of acid precipitation* (eds Drablos, D. & Tolland, A.) 41-46 (SNSF, As-NLH, Norway, 1980).
13. Barrett, C. F. et al. *Acid deposition in the United Kingdom* (Warren Spring Laboratory, Stevenage, 1983).
14. Cape, J. N. *Envir. Pollut.* **A34**, 259-274 (1984).
15. Tukey, H. B. Jr. *A. Rev. Pl. Physiol.* **21**, 305-324 (1970).
16. Hori, T. (ed.) *Studies on Fogs* (Tanne Trading Co., Sapporo, 1953).

Dating Middle Palaeolithic red beds in southern Greece

Kevin O. Pope*, Curtis N. Runnels† & Teh-Lung Ku‡

* Department of Geology, Stanford University, Stanford, California 94305, USA

† Program in Values, Technology, Science and Society, Stanford University, Stanford, California 94305, USA

‡ Department of Geological Sciences, University of Southern California, Los Angeles, California 90089-0741, USA

Late Quaternary red beds with associated Mousterian tools are widely distributed in Greece but have not been previously dated by direct methods. Here we extend the findings of a geological and archaeological study of the interaction between human settlement and landscape in the southern Argolid (Fig. 1) during the late Quaternary^{1,3}. Uranium-series dates were obtained from groundwater carbonate crusts, formed on Mousterian tools found in the southern Argolid, and from pedogenic carbonates found in the red beds that contain them. These dates indicate an age of $\sim 50,000$ yr BP for the Mousterian tools and identify three episodes of red bed formation, beginning in the mid-Pleistocene and ending before the last glacial maximum. The radiometric dates are the first for the Mousterian in southern Greece and the first for alluvial red beds in the Mediterranean region.

The southern Argolid landscape was shaped by episodic changes in geomorphic stability³ and changes in coastlines^{4,5}, which influenced the subsistence and settlement strategies of the prehistoric and historic peoples of the region. Destabilization of the landscape, caused by climatic change in the Pleistocene, and human activities in the Holocene³, resulted in brief periods of valley alluviation separated by long periods of stability and soil formation.

The late Quaternary valley fills of the southern Argolid comprise seven temporally distinct alluvial units³, of which five are exposed in a stream cut at B27, an archaeological site stratified between two red bed deposits (Fig. 2). The oldest unit at B27 is the middle member Loutro alluvium, which is composed of debris flow, braided stream, streamflood and distal fan deposits, although only the distal fan deposits are exposed in the cut. B27 is incorporated in a palaeosol formed in the middle member Loutro alluvium. The site is buried by the upper member Loutro alluvium, which consists mainly of braided stream and streamflood deposits. The palaeosol, besides being well developed, contains grey-green mottles and groundwater carbonate crusts on pebbles and stone tools, indicating a fluctuating water table, perhaps caused by former seasonal springs or aggradation of the nearby stream.

The site consists of a 1-m thick layer of red soil containing chert flakes and retouched tools. It extends laterally for >5 m in the stream bank, but its total extent is not known. No features or bones were noted in the exposure; chert flakes and tools were the only artefacts collected. The majority (70%) of the 218 artefacts collected are unmodified flakes; complete flakes comprise 19% of the collection, and tools make up 8%. All the specimens are small scale, but may be described broadly as Mousterian (Fig. 3). The site is primary: the artefacts, though covered with a carbonate crust, are in very good condition, with sharp, unrolled edges. The very small flakes collected matched closely the size, shape and material of the retouch flake scars on the finished tools, and resulted from tool manufacture.

Other Middle Palaeolithic material in the southern Argolid is known from chance finds⁶, and from the discovery of other surface sites in systematic surveys². One of these sites, B85, produced small Mousterian tools different from those at B27 (Fig. 3).

Middle Palaeolithic sites from Greece (Fig. 1) include those in Epirus, Corfu⁷⁻¹⁰, Elis¹¹ and Larisa (Thessaly)^{12,13}. The artefacts from the Argolid are broadly comparable in typology

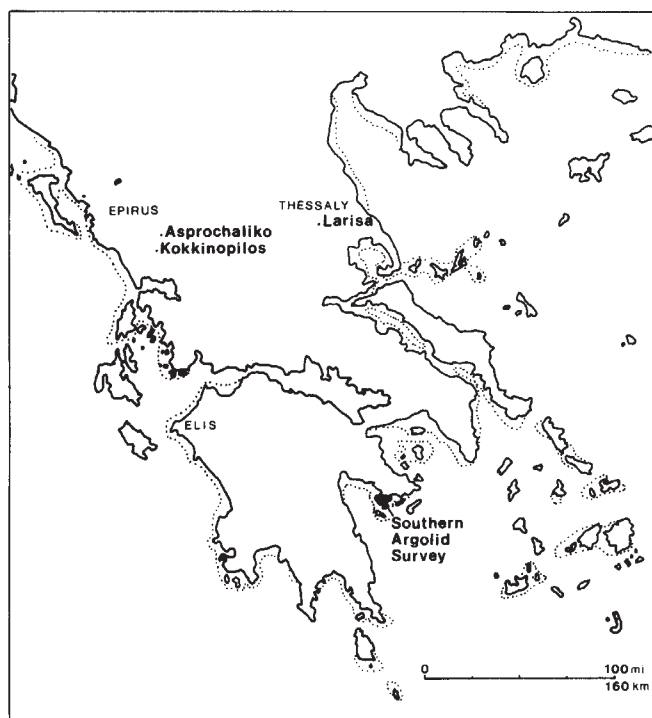


Fig. 1 Map of Greece with archaeological sites and regions mentioned in the text. The broken line represents the estimated location of the coastline at 50,000 yr BP.

to artefacts from these areas. In Epirus, the Middle Palaeolithic was excavated at Kokkinopilos and Asprochaliko⁷. A stratigraphic sequence, similar to that of B27, was found at Kokkinopilos, with Mousterian tools stratified between two red bed deposits and associated with an old land surface. At Asprochaliko, a stratified succession of Palaeolithic industries was found with a typical Mousterian at the bottom, followed by a small Mousterian, and overlain by an Upper Palaeolithic industry. Higgs^{7,8} considered the Kokkinopilos Middle Palaeolithic, with its foliate points, as possibly contemporary to the basal Mousterian at Asprochaliko, but the Epirus data do not preclude their being of a different age. The foliate points at B85 are similar to the Kokkinopilos points and the small Mousterian of B27 is similar in size and typology to the small Mousterian of Asprochaliko. Both B27 and B85 lie on top of the middle member Loutro, but their relative ages are unknown. At present the sequence of Greek Middle Palaeolithic tool industries is that of a typical Mousterian followed by a small Mousterian.

The only previously published dates for these industries are radiocarbon dates of $24,900 \pm 1,100$, $37,000^{+4,100}_{-2,100}$ and $>39,900$ yr BP for the small Mousterian at Asprochaliko¹⁴. That B27 is as old as this is based on typological grounds: it may be placed with the late Mousterian of the Balkan Peninsula and the Near East, dated by a series of radiocarbon dates to between

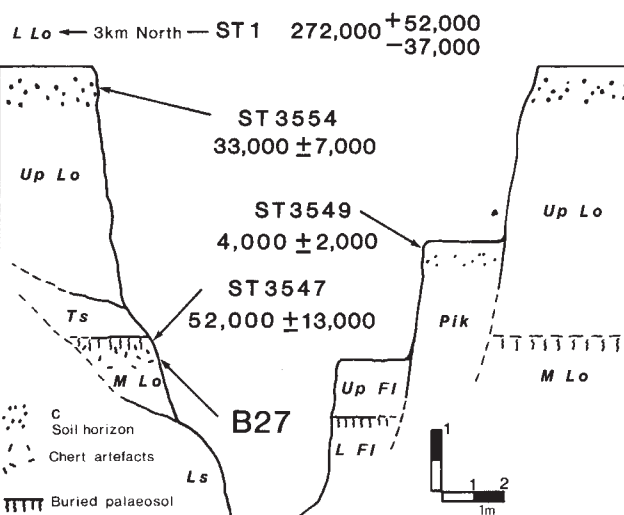


Fig. 2 Composite diagram of alluvial and soil stratigraphy at B27 with locations of carbonate samples used in dating. Ages are in years BP. Abbreviations: Up Lo, upper member Loutro alluvium; M Lo, middle member Loutro alluvium; Pik, Pikrodhafni alluvium; Up Fl, upper member Flamboura alluvium; L Fl, lower member Flamboura alluvium; Ts, talus (derived from limestone); Ls, limestone.

40,000 and 50,000 yr BP¹⁵⁻¹⁷. Existing dates for Greek red beds are based on indirect evidence from global climatic or glacial histories¹⁸, or associations with raised beaches¹⁹, all of which are controversial. The radiocarbon dates from Asprochaliko confirm the Middle Palaeolithic placement of the small Mousterian, but do not provide a date for the red beds outside the cave.

The artefacts recovered from B27 in the Argolid were coated with a groundwater carbonate rind. Pedogenic carbonate nodules are present in the C soil horizons in the upper member Loutro and Pikrodhafni alluvium, near B27, and in the lower member Loutro, 3 km north of the site. These carbonates were dated (Table 1) with uranium-series isotopes, using the techniques proposed by Rosholt²⁰ and later refined by Szabo and Sterr²¹, and Ku and Liang²².

Ku and Liang²² discuss the theoretical basis for the calculation of the carbonate $^{234}\text{U}/^{238}\text{U}$ and $^{230}\text{Th}/^{234}\text{U}$ ratios. The calculation involves two assumptions: (1) the solution from which carbonates form co-precipitates some uranium but no thorium; (2) the acid-leaching procedure adopted does not cause isotopic fractionation. Whereas the second assumption has been verified experimentally²⁰⁻²², the validity of the first assumption must be inferred from general knowledge of the solution chemistry of uranium and thorium. In natural waters, uranium is relatively soluble; in contrast, the easy hydrolysis of Th^{4+} ions renders the thorium isotopes highly insoluble.

The uranium-series dates determined by the above methods for the southern Argolid samples are stratigraphically consistent,

Table 1 Chemical analyses and age determinations of carbonate samples

Sample no.	f_L	Activity ratios in leachate			Activity ratio in residue			Age (kyr)
		$^{230}\text{Th}/^{232}\text{Th}$	$^{234}\text{U}/^{232}\text{Th}$	$^{238}\text{U}/^{232}\text{Th}$	$^{230}\text{Th}/^{232}\text{Th}$	$^{234}\text{U}/^{232}\text{Th}$	$^{238}\text{U}/^{232}\text{Th}$	
ST3549	0.767	0.73 ± 0.02	1.24 ± 0.02	1.17 ± 0.02	0.70 ± 0.02	0.52 ± 0.01	0.53 ± 0.01	4 ± 2
ST3554	0.629	0.75 ± 0.02	0.97 ± 0.03	0.97 ± 0.03	0.63 ± 0.02	0.50 ± 0.01	0.54 ± 0.01	33 ± 7
ST3547	0.819	1.16 ± 0.04	1.12 ± 0.03	1.05 ± 0.03	0.91 ± 0.03	0.48 ± 0.02	0.62 ± 0.02	52 ± 13
ST1	0.934	1.25 ± 0.03	1.32 ± 0.03	1.26 ± 0.03	0.38 ± 0.02	0.46 ± 0.02	0.70 ± 0.02	272^{+52}_{-37}

The samples were ground into a powder, heated to convert CaCO_3 to CaO , and leached with dilute nitric acid. The non-leachable residues were dissolved with HF/HClO_4 . Radiochemical assay was done on both the leachate and residue fractions, using isotope dilution and α -particle spectrometry. The above procedures are outlined elsewhere²⁰⁻²². The uncertainties quoted are 1 s.d., derived from the counting statistics only. f_L is the weight fraction of leachables. Age was calculated using the methods of refs 20-22. Note that these ages differ slightly from those reported earlier³, owing to a more conservative calculation of the errors involved.

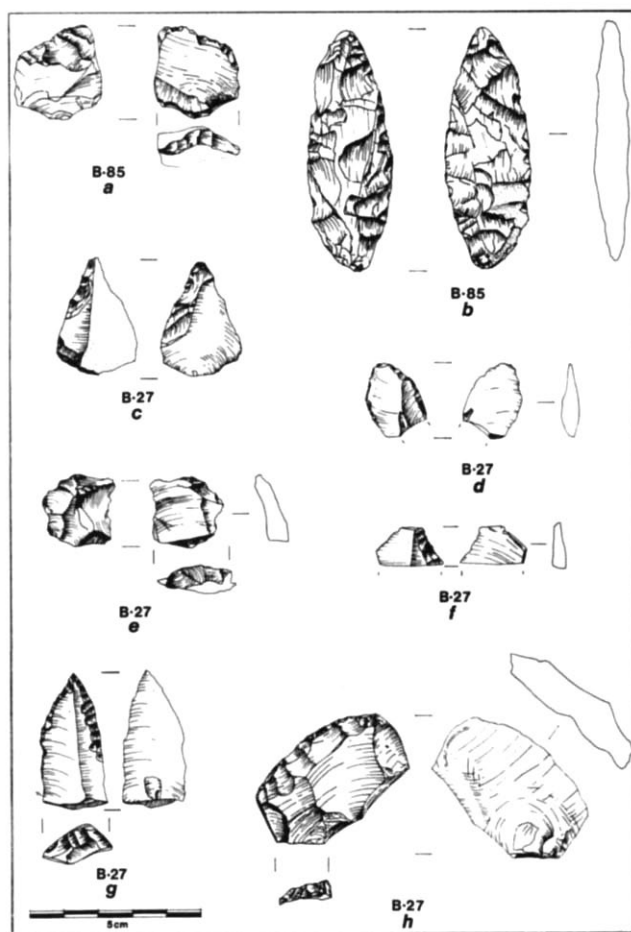


Fig. 3 Small Mousterian tools from B27 and B85 in the southern Argolid: a, B85, flake with faceted platform; b, B85, bifacial foliate point; c, B27, point; d, B27, retouched flake; e, B27, flake with faceted platform; f, B27, retouched flake; g, B27, point; h, B27, side scraper on a Levallois flake.

and agree with the suggested Middle Palaeolithic age of B27. Note that the stone tools are older than groundwater carbonate, but younger than the lower member Loutro. A reasonable minimum age for B27 is 40,000 yr BP, given that time is required for the carbonate crusts to form; the true age is probably closer to 50,000 yr BP.

The pedogenic carbonate dates record when carbonate was formed in the C horizon, which post-dates the deposition of the alluvium. Estimates of the time required for the carbonate nodules to form in semi-arid climates²³ indicate that the abundant hard nodules in the Loutro soils form in a minimum of 9,000 yr and the softer, less frequent nodules in the Pikrodhafni soil form in a minimum of 1,000 yr. Thus we estimate the minimum ages of the lower and upper members of the Loutro alluvium and the Pikrodhafni alluvium as 244,000, 35,000 and 3,000 yr BP, respectively, and the true ages at 281,000, 42,000 and 5,000 yr BP. The last date agrees with the archaeological evidence, as late Bronze Age (~3,500 yr BP) sites are found on top of the Pikrodhafni soil and early Bronze Age sherds (~5,000 yr BP) were deposited with the alluvium³.

The uranium-series dates and red bed stratigraphy indicate an age of 50,000 yr BP for the Middle Palaeolithic industry in the southern Argolid: this is in agreement with the emerging chronology for the eastern Mediterranean¹⁵⁻¹⁷. The small Mousterian at Asprochaliko is typologically similar to the B27 Mousterian and may be older than indicated by radiocarbon determinations¹⁴; the basal Mousterian at Asprochaliko must be older than 50,000 yr BP.

Thus the uranium-series method, as applied in the southern Argolid, seems promising for dating both archaeological sites and late Quaternary red beds. Red bed sequences which contain buried palaeosols, carbonate horizons and Palaeolithic artefacts may prove useful in establishing late Quaternary stratigraphies in the Mediterranean.

We thank T. H. van Andel and members of the Stanford University Archaeological and Environmental Survey of the Argolid for comments and P. Murray for help in preparing the figures. The research was funded in part by the NSF, the US National Endowment for the Humanities, the US National Geographic Society, and the Stanford in Greece Program. Z. C. Liang assisted with the radiometric dating, which was funded by the NSF and the Petroleum Research Fund of the American Chemical Society.

Received 27 August; accepted 12 September 1984.

1. Jameson, M. H. *Ann. N. Y. Acad. Sci.* **268**, 74-91 (1976).
2. Runnels, C. N. *B.A.R. int. Ser.* **155**, 261-264 (1983).
3. Pope, K. O. & van Andel, T. J. *H. J. Archaeol. Sci.* (in the press).
4. van Andel, T. J. H., Jacobsen, T. W., Jolly, J. B. & Lianos, N. *J. Field Archaeol.* **7**, 389-402 (1980).
5. van Andel, T. J. H. & Lianos, N. *Int. J. naut. Archaeol. Underwat. Archeol.* **12A**, 303-324 (1983).
6. Bialor, P. A. & Jameson, M. H. *Am. J. Archaeol.* **66**, 181-182 (1962).
7. Dakaris, S. I., Higgs, E. S. & Hey, R. W. *Proc. prehist. Soc.* **30**, 199-245 (1964).
8. Higgs, E. S. & Vita-Finzi, C. *Proc. prehist. Soc.* **31**, 1-29 (1966).
9. Coles, J. M. & Higgs, E. S. *The Archaeology of Early Man*, 311-322 (Penguin, Harmondsworth, 1969).
10. Sordinas, A. *Balkan Stud.* **10**, 393-424 (1969).
11. Chavaillon, J., Chavaillon, N. & Hours, F. *Bull. Corresp. Hellénique* **91**, 151-201 (1967); **93**, 97-151 (1969).
12. Milošević, V., Boessneck, J., Jung, D. & Schneider, H. *Paläolithikum um Larissa in Thessalien* (Habelt, Bonn, 1965).
13. Theodoridis, D. I. *Avgi tis Thessalikis Proistorias*, 11-34 (Filarkhiou Etairias Volou, Volos, 1967).
14. Bailey, G. N., Carter, P. L., Gamble, C. S. & Higgs, H. P. *Proc. prehist. Soc.* **49**, 15-42 (1983).
15. Smith, F. H. *Curr. Anthropol.* **23**, 667-703 (1982).
16. Jelinek, A. J. *Science* **216**, 1369-1375 (1982).
17. Honea, K. A. *J. Archaeol.* **85**, 483-486 (1981).
18. Dufau, J.-J., Bousquet, B. & Pechoux, P.-Y. *Revue Géol. dynam. Géogr. phys.* **21**, 29-40 (1979).
19. Dufau, J.-J. *Annls. Géogr.* **79**, 325-342 (1970).
20. Rosholt, J. N. *Geol. Soc. Am. Abstr. Progr.* **8**, 1076 (1976).
21. Szabo, B. L. & Sterr, H. *U.S. geol. Surv. Open-File Rep.* **78-701**, 416 (1978).
22. Ku, T. L. & Liang, Z. C. *Nucl. Instrum. Meth. Phys. Res.* **223**, 563-571 (1984).
23. Birkeland, P. W. *Pedology, Weathering, and Geomorphological Research* (Oxford University Press, 1974).

Horizontal cell synapses onto glycine-accumulating interplexiform cells

Robert E. Marc & Wei-Ley S. Liu

University of Texas, Sensory Sciences Center,
6420 Lamar Fleming Blvd, Houston, Texas 77030, USA

Horizontal cells mediate lateral transmission of signals in the outer plexiform layer of the vertebrate retina, and are presumed to contribute to surround properties of photoreceptors and bipolar cells by chemical transmission. The cell bodies¹ and dendrites² of fish horizontal cells possess presynaptic specializations characteristic of conventional chemical synapses. Horizontal cell axon terminals have not so far been shown to contain presynaptic specializations nor have the targets of the somatic and dendritic synapses been fully characterized. Using electron microscope autoradiography of retinas labelled by high-affinity ³H-glycine uptake, we show here that goldfish horizontal cells make somatodendritic and axodendritic synapses on glycinergic interplexiform cells (Gly-IPCs) as apposed to dopaminergic interplexiform cells. Thus, horizontal cells have at least three postsynaptic targets: photoreceptors, bipolar cells and Gly-IPCs. Gly-IPCs may constitute a major alternative pathway for horizontal cell signals to reach the inner plexiform layer.

The use of high-affinity ³H-glycine uptake as an autoradiographic probe has revealed two types of presumed glycinergic

interneurons in fish³⁻⁵ and amphibian retinas^{5,6}: a large population of amacrine cells and a sparse population of interplexiform cells. We refer to the latter as 'Gly-IPCs' to indicate that these neurones are uniquely revealed by high-affinity ³H-glycine uptake while recognizing that their characterization as 'glycinergic' is provisional³⁻⁵. Goldfish Gly-IPCs are situated in the midst of the inner nuclear layer^{3,4}, are the largest cells therein⁴ and possess two to four thick primary dendrites that extend obliquely from the soma, gently inclined towards the outer plexiform layer (Fig. 1a) where they branch further (Fig. 1b). The frequency of Gly-IPCs in horizontal sections of seven sample fields (103 Gly-IPCs and a total area of 3.89 mm²) ranged from 20 to 33 cells mm⁻² (mean 25.9; s.d. 5.4; *n* = 7). The boundaries of the region harbouring Gly-IPCs are indistinct and dividing cell number by area may underestimate frequency. Using mean inter-cell spacing for each sample, we calculated frequencies of 19.8–48.9 cells mm⁻² (the reciprocal square of inter-cell spacing). In any case the cell density is low and comparable with previous estimates^{3,4}.

The true horizontal extent of a single Gly-IPC's arbor is not known, but a partial reconstruction of one cell is shown in Fig. 1c. A single dendritic path can extend at least 200 µm; the arbor diameter could be at least 400 µm and is probably larger. As the processes of Gly-IPCs reach the outer plexiform layer, they become beaded at intervals of 3–5 µm and branch sparingly (Fig. 1b). We cannot prove that all ³H-glycine-labelled terminals in the outer plexiform layer arise from Gly-IPCs, but this seems probable because: (1) a large proportion of them can be traced to Gly-IPC processes in serial sections and (2) conditions which selectively suppress Gly-IPC soma labelling also suppress terminal labelling. The terminal dendrites lie within the narrow plane of the outer plexiform layer, coursing along the distal surfaces of GABAergic (GABA, γ-aminobutyric acid) H1 horizontal cells. Some of the terminals contact and indent the horizontal cell surface and there receive conventional synapses from the GABAergic horizontal cells (Fig. 2a,b). Marc and Lam³ previously reported two examples of such synapses but their morphological preservation was poor. Moreover, we did not know the distributions of such contacts in the outer plexiform layer nor did we suspect their incidence to be so high as to have major physiological implications. We have now examined over 500 well-labelled profiles in the outer plexiform layer and have found 19 distinctive synaptic contacts and many possible contacts. We also found a few unlabelled contacts. As our label criterion is high and horizontal cells are known to be presynaptic to other targets, such as photoreceptor teleodendria², it is likely that the unlabelled profiles represent other recipients of GABAergic horizontal cell synapses. A single Gly-IPC should subtend about 250 GABAergic horizontal cells and, depending on the fraction of overlap with other Gly-IPCs, each cell could receive several hundred synapses. With this large synaptic input, GABAergic horizontal cells could dominate the response properties of Gly-IPCs.

We have also found synaptic output from horizontal cell axon terminals to Gly-IPC dendrites (Fig. 2c, d). We have examined over 100 labelled profiles in the inner nuclear layer and have been able to identify four clear synaptic contacts. The axon terminals of fish horizontal cells have been enigmas: they generate large photoresponses, yet their synaptic targets have never been identified nor have they been shown to contain conventional synapses. Recently, Marshak and Dowling⁷ have also found goldfish horizontal cell axon terminal synapses in conventionally prepared material and, as in the case of somatic synapses, there seem to be several different target cells for the axon terminals. It should be emphasized that the scarcity of horizontal cell synapses may be more apparent than real. First, as we examine these cells with a more practised eye, more synapses become apparent. Second, but more important, Gly-IPCs seem ideally suited to integrating a few synaptic inputs from each horizontal cell or its axon over a large sample of the population, leading to a large net input. Thus, what is rare for the anatomist is abundant for the Gly-IPC.

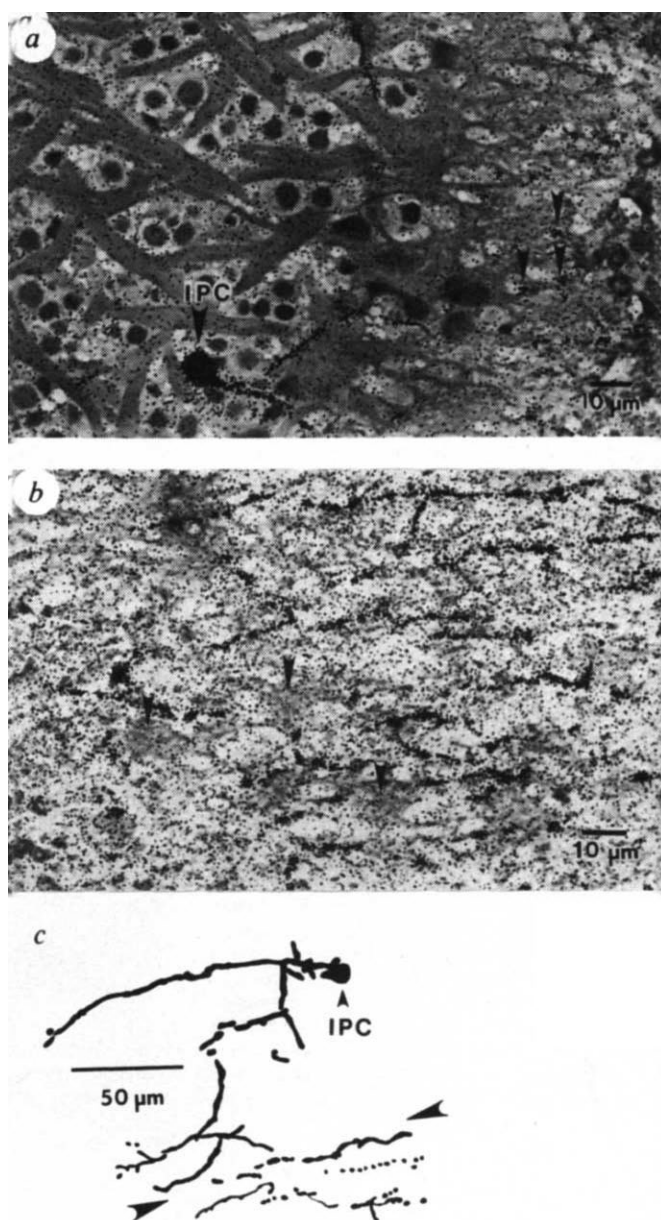
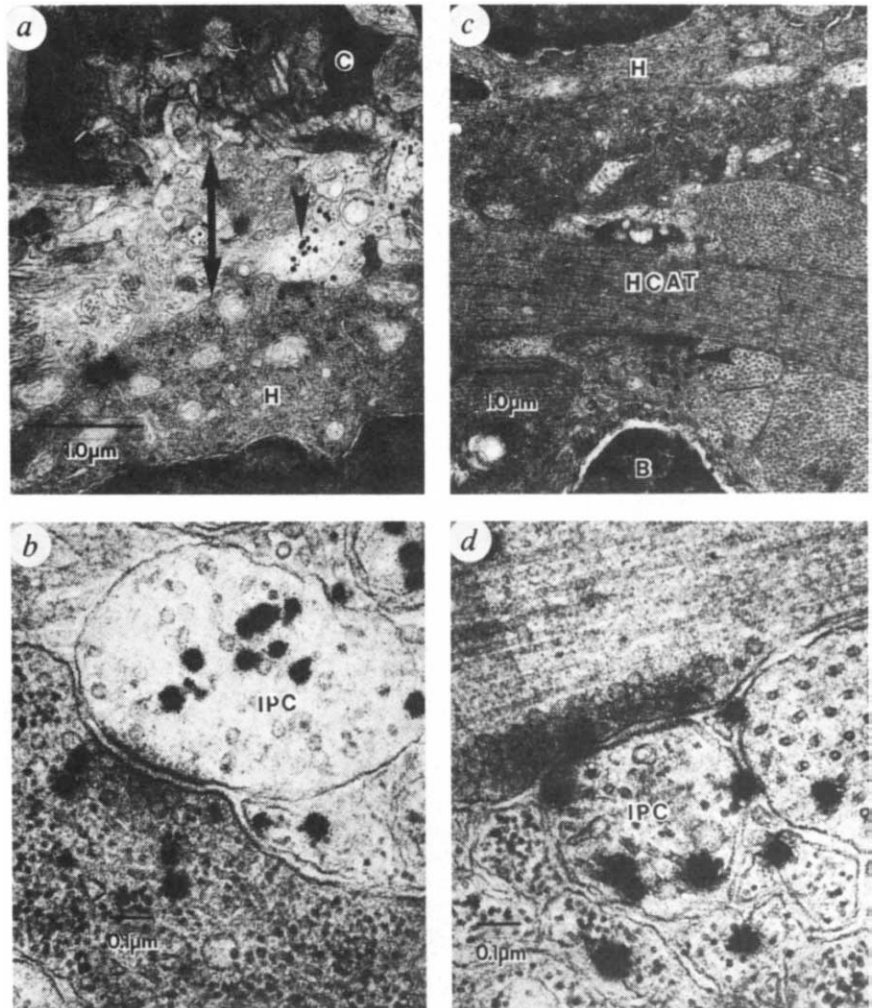


Fig. 1 The morphology of glycine-labelled interplexiform cells (Gly-IPCs) in distal goldfish retina. *a*, Oblique section showing a Gly-IPC in the inner nuclear layer, surrounded by the round nuclei of bipolar cells and the large tuberos horizontal cell axon terminals. This cell has two dendrites coursing towards the outer plexiform layer where there are many labelled terminals, three of which are indicated by arrowheads. *b*, Horizontal section directly through the outer plexiform layer, demonstrating the long, beaded strands of Gly-IPC terminals. The grey blotches indicated by arrows are the most distal fragments of the horizontal cell bodies just bordering the outer plexiform layer. *c*, Partial reconstruction of a Gly-IPC dendritic arbor from nine serial horizontal sections. The cell body gave rise to two identifiable dendrites, one of which was traced into the outer plexiform layer. The plane of border between the inner nuclear layer and the outer plexiform layer is indicated by two arrowheads.

Methods: Isolated goldfish retinas (*Carassius auratus*) were incubated for 10 min in oxygenated physiological saline containing 1–5 µM ³H-glycine at an activity of 100 µCi ml⁻¹, rinsed, fixed and processed for electron microscopy as described previously³. Serial semi-thin sections (0.5–2 µm) were cut for light microscope autoradiography and cell reconstructions. Light microscope autoradiographs were exposed for 3 weeks and developed with Kodak D-19 developer.

Fig. 2 Electron microscope autoradiographs of horizontal cell synapses onto Gly-IPCs. *a*, Low magnification view of a labelled terminal (arrowhead) apposed to the distal surface of a type H1 horizontal cell (H). The labelled terminal lay within the outer plexiform layer proper (OPL, indicated by double-headed arrow) C, cone terminal. *b*, High magnification view of the labelled terminal (IPC) indicated in *a*. Gly-IPC terminals are characteristically pale, occasionally containing scattered glycogen granules and a few small pleomorphic vesicles. The horizontal cell presynaptic specialization consists of a cluster of large round vesicles 50–60 nm in diameter apposed to a 20–25-nm wide synaptic cleft. The horizontal cell ground substance and the vesicles are rather electron-dense. *c*, Low magnification view of a labelled process (arrowhead) apposed to the proximal surface of a horizontal cell axon terminal (HCAT) in the inner nuclear layer. The field includes the cell bodies of a horizontal cell (H) and a bipolar cell (B). Both GABAergic and non-GABAergic horizontal cells have axon terminals and it is not possible to identify the type of axon terminal in this preparation. *d*, High magnification view of the labelled process (IPC) indicated in *c*. Gly-IPC dendrites in the inner nuclear layer are nondescript, contain a few microtubules and typically lack the paleness of their terminal swellings. This process is the recipient of a large synapse from a horizontal cell axon terminal. The vesicles are 50–60 nm in diameter and form a dense cluster within a small platform protruding from the microtubule-filled axon.

Methods: Tissues processed as described in Fig. 1 legend were sectioned at 60–90 nm, floated onto Formvar-coated slot grids, stained with lead citrate and uranyl acetate, carbon-coated and loop-coated with a monolayer of Ilford L4 nuclear emulsion for electron microscope autoradiography. Electron microscope autoradiographs were exposed for 2 months and developed with phenidon. Sections exhibiting low background labelling (<1 grain per 10 μm^2) were scanned thoroughly for every terminal in the outer plexiform layer labelled at a density at least 100 times background (>10 grains per μm^2). The mean label for the 19 somatodendritic and 4 axodendritic synapses was 180 times background.



Biochemical⁸, autoradiographic^{9,10}, immunocytochemical¹¹ and physiological data^{12,13} indicate that H1 horizontal cells are GABAergic. Their primary inputs are red-sensitive cones¹⁴ and their photoresponses are sustained hyperpolarizations^{15–17}. GABA typically mediates a postsynaptic conductance increase to chloride ions^{18–20}. If the GABAergic horizontal cell synapse behaves as most central nervous system GABAergic synapses, hyperpolarization of H1 horizontal cells should cause a decrease in the rate of GABA release, reducing the magnitude of a dark-maintained inhibitory postsynaptic potential in the Gly-IPCs, yielding a sustained depolarization. The Gly-IPC should possess a spectrum of physiological qualities (except for response polarity) similar to those of H1 horizontal cells: sustained responses, extensive spatial integration, linearity and low-pass filter properties²¹, and dominance by red-sensitive cones. This does not mean that putative inputs from the inner plexiform layer cannot influence the Gly-IPC, but rather that the horizontal cell input seems so large that it cannot be ignored in assessing the probable functions of Gly-IPCs.

The Gly-IPC is a possible pathway for the signals of retinal horizontal cells to reach the inner plexiform layer without passing through bipolar cells and is the first neurochemically identified target of both somatic and axonal horizontal cell synapses. The significance of this cannot be evaluated fully without knowing the connections of these cells in the inner plexiform layer, but such a pathway was predicted by Naka *et al.*²² from physio-

logical data obtained in the catfish retina. Unfortunately, ³H-glycine-accumulating amacrine cells outnumber Gly-IPCs by nearly 100 to 1 (ref. 4), and it is currently impossible to discriminate between interplexiform and amacrine synapses in the inner plexiform layer. The Gly-IPC process connecting the outer and inner plexiform layers is a stout fibre extending from the outer plexiform layer to at least the middle of the inner plexiform layer^{3–5} and there should be no electrotonic constraints on horizontal cell signals reaching the inner plexiform layer via Gly-IPCs.

Our results provide evidence for a second class of interplexiform cell in addition to the well-known dopaminergic interplexiform cell²³. Dopaminergic interplexiform cells were the first known non-photoreceptor inputs to horizontal cells and, paradoxically, Gly-IPCs are the first identified recipients of horizontal cell synapses not associated with the photoreceptor-bipolar cell chain in the outer retina. This suggests that the transfer of surround information to the inner plexiform layer through bipolar cells is but a part of the repertoire of the horizontal cell and that interplexiform cells are critical intermediaries in a wider spectrum of horizontal cell functions. There are now three known targets of horizontal cell chemical synapses: photoreceptor cells², bipolar cells^{24,25} and ³H-glycine-accumulating interplexiform cells.

This research was supported by NIH grant EY02576 (R.E.M.) and NIH Biomedical Research Support Grant NS19342.

Received 27 May; accepted 4 September 1984.

- Witkovsky, P. & Dowling, J. E. *Z. Zellforsch.* **100**, 60-82 (1969).
- Sakai, H. & Naka, K.-I. *Vision Res.* **23**, 339-351 (1983).
- Marc, R. E. & Lam, D. M. K. *J. Neurosci.* **1**, 152-165 (1981).
- Marc, R. E. *Vision Res.* **22**, 589-608 (1982).
- Marc, R. E. in *Retinal Transmitters and Modulators: Models for the Brain* Vol. 1 (ed. Morgan, W. W.) (CRC Press, Boca Raton, Florida, in the press).
- Rayborn, M. E., Sarthy, P. V., Lam, D. M. K. & Hollyfield, J. G. *J. comp. Neurol.* **195**, 585-593 (1981).
- Marshak, D. & Dowling, J. *Soc. Neurosci. Abstr.* **10**, 21 (1984).
- Lam, D. M. K. *Nature* **254**, 345-347 (1975).
- Lam, D. M. K. & Steinman, L. *Proc. natn. Acad. Sci. U.S.A.* **68**, 2777-2781 (1971).
- Marc, R. E., Stell, W. K., Bok, D. & Lam, D. J. K. *J. comp. Neurol.* **182**, 221-245 (1978).
- Lam, D. M. K. *et al. Nature* **278**, 565-567 (1979).
- Murakami, M., Shimoda, Y. & Nakatani, K. *Sensory Processes* **2**, 334-338 (1978).
- Murakami, M., Shimoda, Y., Nakatani, K., Miyashi, E.-I. & Watanabe, S.-I. *Jap. J. Physiol.* **32**, 911-926 (1982).
- Stell, W. K. & Lightfoot, D. O. *J. comp. Neurol.* **159**, 473-502 (1975).
- Kaneko, A. *J. Physiol., Lond.* **207**, 623-633 (1970).
- Mitarai, G., Asano, T. & Miyake, Y. *Jap. J. Ophthalmol.* **18**, 161-176 (1974).
- Hashimoto, Y., Kata, A., Inokuchi, M. & Watanabe, K. *Vision Res.* **16**, 25-30 (1976).
- Curtis, D. R. & Johnston, G. A. R. *Ergebn. Physiol.* **69**, 97-188 (1974).
- Matthews, G. & Wickelgren, W. O. *J. Physiol., Lond.* **293**, 393-415 (1979).
- Miller, R. F., Frumkes, T. E., Slaughter, M. & Dacheux, R. F. *J. Neurophysiol.* **45**, 743-763 (1981).
- Naka, K.-I., Marmarelis, P. Z. & Chan, R. Y. *J. Neurophysiol.* **38**, 92-131 (1975).
- Davis, W. & Naka, K.-I. *J. Neurophysiol.* **43**, 807-831 (1980).
- Dowling, J. E. & Eninger, B. *Proc. R. Soc. B207*, 7-26 (1978).
- Dowling, J. E. & Werblin, F. S. *J. Neurophysiol.* **32**, 315-338 (1969).
- Dowling, J. E., Brown, J. E. & Major, D. *Science* **153**, 1639-1641 (1966).

Homology of Ti α -subunit of a T-cell antigen-MHC receptor with immunoglobulin

Marina Fabbi*†, Oreste Acuto*, John E. Smart† & Ellis L. Reinherz*

* Division of Tumor Immunology, Dana-Farber Cancer Institute and Departments of Medicine and Pathology, Harvard Medical School, Boston, Massachusetts 02115, USA

† Biogen Research Corporation, Cambridge, Massachusetts 02142, USA

Human T-cell receptors for antigen and major histocompatibility complex (MHC) determinants have now been defined on inducer¹, suppressor², and class 1 and class 2 MHC-specific cytotoxic T lymphocytes³ as T3-associated clonotypic molecules (Ti) of relative molecular mass 90,000 (90K) composed of one 49-54K α - and one 43K β -subunit which are disulphide-linked⁴. In the case of the Ti β -subunit, N-terminal amino acid sequencing⁵ and molecular cloning techniques^{6,7} led recently to identification of the Ti β -gene and showed that T-specific V, D, J and C segments fuse to form an active β -gene^{8,9}. So far, however, there have been little structural data available on the Ti α -subunit. Here we have derived the amino acid sequence of a portion of the Ti α -subunit by CNBr fragmentation. Sequence analysis reveals ~40% homology between the Ti α -subunit fragment and the third framework of the variable region of immunoglobulin light and heavy chains, supporting the notion that the Ti α -subunit is a member of the immunoglobulin-Ti β -gene family.

Although the β -subunit of the T-cell receptor for antigen-MHC has been characterized extensively in both mouse and man^{10,11}, little is known about the Ti α -subunit. Nevertheless, comparative two-dimensional peptide mapping analysis of separated Ti α and β molecules isolated with non-cross-reactive anti-clonotypic monoclonal antibodies from human T-lymphocyte clones demonstrated the presence of variable and constant region domains in the α as well as β molecules¹². In the case of the β -subunit, the structural basis for this variability has now been defined with the elucidation of its V, D, J and C regions. These findings imply that the diversity of the T-cell receptor is, at least in part, generated by modular elements rearranging in a manner analogous to immunoglobulin genes in B lymphocytes¹³.

† Present address: Laboratorio di Immunobiologia, Istituto Nazionale per la Ricerca sul Cancro, Viale Benedetto XV, 10, 16132 Genova, Italy.

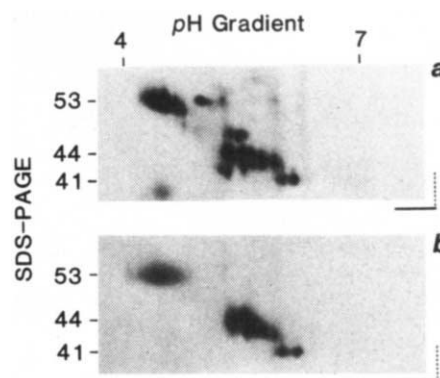


Fig. 1 Purification of the REX Ti molecule. *a*, REX Ti affinity-purified on anti-Ti_{3A} antibody column. *b*, Same preparation after absorption on a wheat-germ lectin column.

Methods: Crude membranes from REX tumour T-cell lines were prepared, solubilized and the lysate subjected to affinity chromatography using anti-Ti_{3A} antibody (IgG1) as described previously⁵. The eluate was neutralized with 0.5 M NaH₂PO₄ and brought to pH 8.0 with 1 M Tris-HCl, then incubated for 60 min with Protein A-Sepharose (Pharmacia) and absorbed subsequently on a chromatographic column of lectin from *Triticum vulgaris* (wheat germ) insolubilized on cross-linked beaded agarose (Sigma); 20 mg of lectin in a packed gel volume of 1 ml were used for lysate equivalent to 5 × 10¹⁰ cells. Absorption was performed by recirculating the sample through the lectin column with a peristaltic pump at a flow rate of 5 ml h⁻¹, overnight at 4 °C. After washing with 100 ml of 20 mM Tris-HCl, 0.5% (w/v) deoxycholate (DOC), 0.02% (w/v) SDS 0.5 M NaCl pH 8, the absorbed material was eluted with 0.5 M N-acetylglucosamine (Sigma) in 20 mM Tris-HCl, 0.5% DOC, 0.15 M NaCl, buffer. The entire procedure was monitored using small aliquots (100 µl) of purified material ¹²⁵I-labelled by the chloramine-T method as tracer. Radioactive fractions were collected and dialysed for 16 h at 4 °C against 0.5% DOC pH 8.0. The purity of the preparation was tested by two-dimensional gel electrophoresis (O'Farrell) of ¹²⁵I-labelled aliquots as described elsewhere⁵.

To obtain data on the primary structure of the Ti α -chain, Ti molecules were isolated from a human tumour line of T-cell origin (REX)⁴. Crude membranes of REX were solubilized and Ti molecules purified by affinity chromatography using anti-clonotypic monoclonal antibody (anti-Ti_{3A}) with a method described previously⁵. To assess the purity of the material obtained, a small aliquot was labelled with ¹²⁵I by the chloramine-T method¹⁴ and analysed by two-dimensional gel electrophoresis¹⁵. Consistent with earlier findings, three major proteins were detected (Fig. 1*a*): a glycosylated 53K α -chain (pI 4.5), a glycosylated 44K β -chain (pI 6) and an intracytoplasmic 41K β -chain (pI 6.5)⁵. In addition, two major contaminants having a relative molecular mass (*M_r*) identical to that of the α -chain were seen (Fig. 1*a*). To eliminate the latter, an additional purification step consisting of a lectin column was used. Previous work in our laboratory has demonstrated that the Ti molecule binds specifically to concanavalin A, lentil lectin and wheat-germ agglutinin; for our purification we chose wheat-germ agglutinin as it most readily retained the Ti molecule and eliminated contaminating material (Fig. 1*b*).

Given this degree of purity, the α - and β -subunits were separated by preparative gel electrophoresis in reducing conditions; the material corresponding to the Ti α band was electroeluted and subjected to NH₂-terminal amino acid sequencing using a solid-liquid-gas phase sequenator. However, attempts to obtain NH₂-terminal amino acid sequence data gave negative results, indicating that the amino-terminal residue was blocked. In this respect, it is known that many proteins, including some immunoglobulin chains, have blocked amino-terminal residues, commonly through acetylation or cyclization of an amino-terminal glutamyl residue to pyrrolidonecarboxyl (pyroglutamyl) residues, or alternatively, via modification of the first

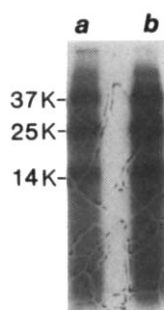


Fig. 2 CNBr cleavage of the REX Ti α -subunit. Lanes *a* and *b* show two different preparations. The purified material eluted from the lectin column and dialysed was lyophilized, resuspended in cold (-20°C) absolute ethanol and left overnight at -20°C to precipitate proteins. After centrifugation, the insoluble pellet was dried under vacuum, resuspended in Laemmli reducing sample buffer²² with an aliquot of ^{125}I -labelled purified protein, then subjected to electrophoresis on a 10% acrylamide preparative SDS gel with 0.1 mM thioglycolate (Sigma) in the reservoir buffer. The gel was then immediately exposed for autoradiography and the radioactive bands corresponding to the α - and β -subunits were excised and electroeluted in 50 mM sodium bicarbonate, 0.1% SDS buffer. Quantitation by silver nitrate staining²³ of an aliquot of the eluted material showed that ~ 1 nmol of purified Ti was obtained from 3×10^{11} REX cells. The electroeluted Ti α -chain was lyophilized, redissolved in 100% H_2O containing 5% β -mercaptoethanol, boiled for 1 min, then re-lyophilized. The sample was then dissolved in 500 μl of 70% formic acid containing 100 mg CNBr, cleaved for 18 h at room temperature in the dark, then diluted to 7 ml with H_2O and lyophilized. Subsequently, CNBr fragments were dissolved in H_2O and a 2% aliquot relabelled with ^{125}I to be used as tracer for the unlabelled material. The cleaved α -subunit was subjected to electrophoresis on an 18% acrylamide preparative SDS gel, the gel was autoradiographed while wet and the ^{125}I bands excised for electroelution. After electroelution, the fragments were dialysed extensively against 0.1% SDS, lyophilized and then subjected to sequential Edman degradation on a gas-liquid-solid phase protein sequencer (Applied Biosystem model 470A).

amino acid to *N*-formyl methionine or *N*-trimethylalanine¹⁶. The β -subunit from the same preparations permitted unambiguous NH_2 terminal amino acid analysis.

To circumvent this problem, we performed CNBr cleavage of the separated α -chain, then an aliquot of this preparation was labelled with ^{125}I by the chloramine-T method and analysed on an 18% acrylamide gel. Figure 2 shows that the intact Ti α -subunit was virtually undetectable. Instead, three major fragments with apparent relative molecular masses of 37K, 25K and 14K, respectively, were detected. A small aliquot of this labelled material was re-iodinated and used as tracer for electrophoretic separation of the CNBr peptides on an 18% acrylamide preparative gel. The bands corresponding to each of the molecular species were isolated and the fragments were electroeluted individually for amino acid sequencing. Of the three fragments obtained, only the 25K fragment yielded sequence information; unambiguous amino acid residue assignments could be made from 16 of 19 cycles. Of the three residues remaining, positions 13 and 16 were scored as probably aspartic acid and arginine, respectively; position 18 could not be assigned. In contrast, the 37K and 14K peptides appeared to be blocked. Figure 3 shows the amino acid sequence of the 25K CNBr fragment of the Ti α -chain. Note that a methionine has been placed in the first position, as CNBr cleavage occurs at the methionyl peptide bonds. Although tryptophan can also act as a cleavage site for CNBr, this is unlikely as the protein was reduced immediately before CNBr cleavage. In addition, high concentrations of CNBr (200 mg ml^{-1} in 70% formic acid) as used here yield negligible amounts of material from tryptophan¹⁷. Moreover, the fragment was not generated from an acid cleavage (Asp-Pro bond)



Fig. 3 Amino acid sequence determination of REX Ti α CNBR fragments. The amino acid sequences are given in the single-letter code. The Ti α - and β -subunits are derived from the REX tumour; BS-1 is a rabbit immunoglobulin κ V region, as is BS-5; G1MSAA is a mouse immunoglobulin heavy-chain V region; 93G7 is a mouse immunoglobulin heavy-chain V region. Numbers on the left indicate the position of the residue in the molecule. Numbers above the top line indicate the positions of the assigned residues in the REX Ti α fragments. Shading indicates homology. Amino acid sequence data were obtained by using a gas-liquid-solid phase protein sequencer (model 470A; Applied Biosystems, Foster City, California) essentially as described by Hewick *et al.*²⁴. The resulting phenylthiohydantoin amino acid residues were analysed by reverse-phase HPLC on an IBM Cyano analytical column essentially as described by Hunkapiller and Hood²⁵. Computer search was done using the Protein Sequence Data Base¹⁹ of the National Biomedical Research Foundation (Georgetown University Medical Center).

because tyrosine rather than proline was detected as the first amino acid¹⁸.

To determine whether there were any homologies between known proteins and the REX Ti α sequence, we searched the Dayhoff protein data base¹⁹. Of the 23 best matches, 15 were with variable regions of immunoglobulin heavy and light chains, including rabbit immunoglobulin κ variable (V) region, mouse immunoglobulin heavy-chain V (V_H) region and human, rabbit and murine immunoglobulin λ V region. Figure 3 shows several representative examples. Note that the REX α -chain sequence shows 42–47% similarity to a stretch of amino acids in framework 3 of the immunoglobulin V regions. Specifically, the homology involves the amino acids around and including the cysteine at about position 100, which corresponds to the second residue forming the intrachain disulphide bond within the variable region. This location is just NH_2 -terminal of the third hypervariable region of immunoglobulin and is the most highly conserved sequence of immunoglobulin variable regions²⁰.

Given the homology of the Ti α sequence with immunoglobulin V_H regions, it was important to exclude the possibility that the sequence was derived from the heavy chain of the anti-Ti_{3A} antibody used in affinity purification. Therefore, we determined the pI of the anti-Ti_{3A} heavy chain (7.0) and demonstrated that such a molecule was absent from the eluted material (Fig. 1 and data not shown). In addition, antiserum raised against the purified Ti α molecule used for sequencing did not react with human B lymphocytes (M. F. *et al.*, in preparation). Moreover, Fig. 3 shows that the REX Ti α sequence is 36% homologous with the Ti β -chain sequence obtained from the same tumour²¹. Again, this involves residues around the cysteine at position 90 which forms a putative intrachain disulphide bond to the cysteine at position 21 within the Ti β V region.

Based on the location of the above homologies in immunoglobulin and Ti β proteins, we predict that the Ti α -subunit sequence within the 25K fragment is derived from a homologous position in the V gene of the Ti α -subunit. This is also consistent with the fact that the 25K fragment may result from cleavage of the NH_2 -terminally blocked 37K fragment thus lying approximately 100 amino acid residues downstream of the N-terminus. That the 14K CNBr fragment was also NH_2 -terminally blocked further supports this view. Moreover, any

discrepancies between this predicted number of amino acids and the size of the latter fragment may be due to inaccuracy of molecular weight determination for small peptides, which may be glycosylated, in SDS gels. Note that the fragments were localized by their ability to be iodinated, therefore any additional fragments may have escaped detection.

As noted previously, peptide mapping analysis provided evidence of variability at the protein level in Ti α -subunits derived from clones of differing specificities. The homology between the REX Ti α sequence and the sequences of immunoglobulin heavy, κ and λ chains and the Ti β V region around the central cysteine (residues 80–100) demonstrated here suggests that a structurally similar V-region domain is contained

within the Ti α -subunit also. Furthermore, the presence of an immunoglobulin-like V region in the Ti α - and β -subunits suggests that both chains are required for antigen and/or MHC recognition, as is the case for immunoglobulin heavy and light chains. In this respect it is interesting to speculate that the fundamental structural basis of T- and B-cell recognition may be similar. In any event, these results support the notion that there exists a gene superfamily encoding B- and T-cell receptor molecules, including the Ti α -subunit.

This work was supported in part by NIH grants RO1 AI 19801, AI 21226 and CA 19589. We thank Ms Selma Oppenheim for her sterile culture work with the REX cell line and Ms Tanya Fabel for assistance with the computer search.

Received 17 September; accepted 8 October 1984.

1. Meuer, S. C. *et al. Science* (1983) **222**, 1239–1242 (1983).
2. Bensussan, A., Acuto, O., Hussey, R. E., Milanes, C. & Reinherz, E. L. *Nature* **311**, 565–567 (1984).
3. Meuer, S. C. *et al. Nature* **303**, 808–810 (1983).
4. Acuto, O. *et al. Cell* **34**, 717–726 (1983).
5. Acuto, O. *et al. Proc. natn. Acad. Sci. U.S.A.* **81**, 3851–3855 (1984).
6. Yanagi, Y. *et al. Nature* **308**, 145–149 (1984).
7. Hendrick, S. M., Cohen, D. I., Nielson, E. A. & Davis, M. M. *Nature* **308**, 145–149 (1984).
8. Siu, G. *et al. Cell* **37**, 393–401 (1984).
9. Kawaler, J., Davis, M. M. & Chien, Y. *Nature* **310**, 421–423 (1984).
10. Kaffler, J., Kubo, R., Haskins, K., White, J. & Marrack, P. *Cell* **34**, 727–737 (1983).
11. McIntyre, B. W. & Allison, J. P. *Cell* **34**, 739–746 (1983).
12. Acuto, O., Meuer, S. C., Hodgdon, J. C., Schlossman, S. F. & Reinherz, E. L. *J. exp. Med.* **158**, 1368–1373 (1983).

13. Tonegawa, S. *Nature* **302**, 575–581 (1983).
14. Greenwood, F. C., Hunter, W. M. & Glören, J. S. *J. Biochem. Tokyo*, **89**, 114–116 (1963).
15. O'Farrell, P. H. *J. biol. Chem.* **250**, 4007–4021 (1975).
16. Allen, G. *Sequencing of Proteins and Peptides* Elsevier Biomedical, Amsterdam, (1981).
17. Ozols, J. & Gerard, C. *J. biol. Chem.* **252**, 5986–5989 (1977).
18. Landon, M. *Meth. Enzym.* **47**, 145 (1977).
19. Dayhoff, M. O. *et al. Protein Sequence Data Base: Computer System NBR Rep.* 08710–811001 (National Biomedical Research Foundation, Washington, D.C., 1984).
20. Kabat, E. A., Wu, T. T., Bilofsky, H., Reid-Miller, M. & Perry, H. *Sequences of Proteins of Immunological Interest* (U.S. Dept of Health and Human Services, 1983).
21. Royer, H. D. *et al. Cell* (in the press).
22. Laemmli, U. K. *Nature* **227**, 680–685 (1970).
23. Merrill, C. R., Goldman, D., Sedman, S. A. & Ebert, M. A. *Science* **211**, 1437–1439 (1981).
24. Hewick, R. M., Hunkapiller, M. W., Hood, L. E. & Dreyer, W. J. *J. biol. Chem.* **256**, 7990–7997 (1981).
25. Hunkapiller, M. W. & Hood, L. E. *Meth. Enzym.* **91**, 484–490 (1983).

Somatic mutation and the maturation of immune response to 2-phenyl oxazolone

Gillian M. Griffiths, Claudia Berek, Matti Kaartinen* & Cesar Milstein

MRC Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, UK

Studies on the development of the immune response suggest that the repertoire of expressed antibody specificities is strongly influenced by antigen (reviewed in ref. 1). One way in which this influence is manifested is by a progressive increase in the affinity of antibody for antigen with time after immunization. This phenomenon, termed the 'maturation' of the immune response, must be due to a change in the structure of the antibody being synthesized. However, the precise nature of the changes involved and the genetic mechanisms used to produce them have not been clearly defined. We have now investigated the maturation of the immune response to the hapten 2-phenyloxazolone by mRNA sequencing of specific hybridomas. We conclude that somatic mutation of germ-line encoded genes plays a major role in the generation of antibodies with increased affinity for oxazolone with time after immunization.

The formation of a functional V region requires the integration of three sets of heavy chain (V_H , D and J_H) and two sets of light chain (V_L and J_L) germ-line gene segments. Antibody diversity can arise in several different ways. First, germ-line diversity is provided by the existence of multiple copies of each of these gene segments (in the mouse there are four functional J and 100–300 V segments for each chain and at least 12 D segments). Second, combinatorial diversity arises as a result of the use of different combinations of V , (D) and J and different pairings of light and heavy chains. Third, junctional diversity results from variation in the precise site of joining of the different gene segments. Finally, it is now generally accepted that somatic

point mutation can cause amino acid substitutions throughout the V region, increasing the number of available structures (reviewed in ref. 2).

To learn more about the role that such mechanisms play in producing an effective antibody repertoire during an immune response, we studied V gene expression at several different stages in the response to a simple antigen, 2-phenyl oxazolone. Hybridomas expressing anti-oxazolone antibodies were generated at three different stages after immunization and their V regions analysed by direct sequencing of the immunoglobulin mRNA. The analysis of hybridomas produced at seven days after primary immunization with oxazolone³ demonstrates that the majority of hybridomas express identical V_H and V_L sequences, presumably as the result of expression of a single pair of unmutated germ-line genes which we have termed V_H -Oxl and V_L -Oxl respectively. The V_L -Oxl gene has now been identified and sequenced from a genomic BALB/c library and shows absolute identity with the predicted germ-line sequence (J. Even and C.M., manuscript in preparation).

At seven days after primary immunization, 11 out of 15 hybridomas (73%) express V_H -Oxl and V_L -Oxl. By 14 days, six out of 11 (54%) expressed this combination and after secondary immunization the frequency was four out of 22 (18%) (Table 1). Therefore hybridomas expressing the Oxl sequences are found at all stages of the response analysed. They occur more frequently amongst hybridomas derived after primary immunization.

The Oxl sequences expressed are highly conserved throughout the response and the preferred combinations of V, D and J observed at day seven³ are also seen at later stages. For example in the heavy chain, V_H -Oxl is joined consistently to a D segment of apparently three amino acids in length and a J_H segment which is shorter than its germ-line counterpart. The overall length of the heavy chain variable region is always conserved (Fig. 1). In the light chain there is a strong preference for the use of J_K5 and an absolute conservation of a leucine residue at position 96 (Fig. 2), a position where sequence variation is often introduced during V-J recombination. Thus the combinatorial and junctional mechanisms which can generate antibody diversity are expressed with the same characteristic pattern at all stages of the response. These restrictions conserve certain features of the amino acid sequence which are presumably

* Present address: Department of Bacteriology & Immunology, University of Helsinki, Haartmaninkatu 3, 00290 Helsinki 29, Finland

Table 1 The frequency of the V_H -Ox1 and V_K -Ox1 basic sequences amongst anti-phOx secreting hybridomas derived at different stages of immunization

	LIGHT CHAIN		HEAVY CHAIN		TOTAL NUMBER		
	V_K -Ox-1	Others	V_H -Ox-1	Others	V_H -Ox ₁ - V_K -Ox ₁ combination	Other combinations	% V_H -Ox ₁ - V_K -Ox ₁
DAY 7	NQ2-12.4 -16.2 -17.4 -20.5 -48.2 NQ5-4.3 -61.1 -96.2 -78.2 NQ6-8.3 -24.2	NQ2-45.1 -6.1 -45.10.4 NQ5-89.4	NQ2-12.4 -16.2 -17.4 -20.5 -48.2 NQ5-4.3 -61.1 -96.2 -78.2 -89.4 NQ6-8.3 -24.2	NQ2-45.1 -6.1 -45.10.4	11	4	73%
DAY 14	NQ7-3.3 -34.3 -1.3 -41.3 -5.3 -24.6 -51.3 -6.1.1	NQ7-47.1 -7.1 -23.1	NQ7-3.3 -34.3 -1.3 -41.3 -5.3 -24.6	NQ7-51.3 -6.1.1 -47.1 -7.1 -23.1	6	5	54%
SECONDARY	NQ10-2.2.5 -12.4.6 -15.9 -12.5 -15.3 NQ11-8.1 -1.14 -7.12	NQ10-12.4.7 -11.1 -2.12.4 -2.12.8 -12.22 -2.22 -2.3 -3.8 -4.6 -1.12 NQ11-5.2 -8.4 -1.18 -7.22	NQ10-2.2.5 -12.4.6 -15.9 -12.4.7 NQ11-1.14	NQ10-12.5 -15.3 -2.12.4 -2.12.8 -12.22 -2.22 -2.3 -1.12 -3.8 -11.1 -4.6 NQ11-5.2 -8.4 -7.22 -1.18 -8.1 -7.12	4	18	18%

The first four columns list hybridomas according to whether they express V_H -Ox1, V_K -Ox1 basic sequences or other heavy and light chain sequences respectively. The total number of hybridomas expressing both V_H -Ox1 and V_K -Ox1 at each stage is given in the fifth column, and expressed as a percentage in the sixth column. The sequence of NQ11-1.14 is not given in Figs 1 and 2 as only partial sequence data are available.

important in maintaining the antigenic specificity of the antibody.

A comparison of hybridomas from the day seven, day 14 and secondary responses demonstrates two important changes in the expression of the Ox-1 antibodies (Fig. 3). First, the V -Ox1 sequences diversify. At day seven the putative V_H -Ox1 and V_K -Ox1 germ-line V genes are expressed in an unmutated form. At day 14 and in the secondary response, however, all hybridomas expressing these same basic sequences show nucleotide changes from the putative germ-line sequences which would produce amino acid substitutions in all of the antibodies studied. Second, the antibodies from the later stages of the response have a much higher affinity for antigen. (Fig. 3). It therefore seems that by producing a small number of changes in the Ox1 germ-line sequences, antibodies with a significantly higher affinity for oxazolone have been generated.

Did these variants arise by somatic point mutation or by expression of a different set of closely related germ-line genes? Somatic mutation can be defined unambiguously in the J segments, as the germ-line counterparts of the J segments expressed by BALB/c mice have been sequenced^{4,5}. Changes which occur near the V - J boundary may have arisen during somatic recombination of the V and J segments. However, other changes in the J segment must have arisen by somatic point mutation. Such changes are observed in three different hybridomas (NQ7-5.3, NQ7-6.1 and NQ10-2.2.5). These changes do not represent germ-line differences in individual mice, because the same J segments

are expressed without these changes in other anti-oxazolone hybridomas derived from the same mice. Thus somatic variants can be unambiguously defined by point mutations in the J segments.

In the V gene segment somatic versus germ-line differences are more difficult to distinguish. Unique V_H genes differing by only two nucleotides have been defined in the mouse germ-line⁶. It is therefore not possible to conclude that V regions differing by very few or even a single nucleotide originate from the same germ-line gene. At day seven the majority of hybridomas expressed a single V_H sequence, presumably derived from a single germ-line gene (V_H -Ox1). However, amongst the 10 variant antibodies described from the day 14 and secondary responses there is not a single example of a repeated sequence. The lack of direct repetition of any given variant sequence argues strongly against a germ-line origin for all of the variants. We conclude that most if not all of the variants arose somatically.

The detailed analysis of the light chain sequences demonstrates that the majority of changes are clustered around residue 34, which appears to be a hotspot for changes in the V_K -Ox1 sequence. In seven V_K -Ox1 light chains the first position of codon 34 is changed from C to A, whereas in four other cases a C to G difference is observed at the third position of the codon (Fig. 2). Both differences cause amino acid substitutions, the histidine residue being replaced with either asparagine or glutamine. Although a single base change in the histidine codon could give rise to seven different amino acid substitutions, the two changes

OXAZOLONE HEAVY CHAINS																			10	20	30	40
VH-Ox1	Q	V	Q	L	K	E	S	G	P	G	L	V	A	P	U	C	Q	P				
Q	CAG	GAG	CAG	CUG	AAU	GAG	UCA	GGA	CCU	GUC	CUG	GUG	GUU	CCC	UCA	CAG	AGC	CCG				
NQ7-3.3	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---				
-34.3	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---				
-1.3	---	B	---	f	f	---	---	---	---	---	---	---	---	---	---	---	---	---				
-41.3	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---				
-5.3	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---				
-24.6	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---				
-51.3	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---				
-6.1	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---				
NQ10-2.2.5	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---				
-12.4.6	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---				
-15.9	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---				
-12.4.7	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---				

																			50	60	70	80
VH-Ox1	P	G	K	G	L	E	W	L	G	V	A	P	U	C	Q	P	U	L				
P	CCA	GGA	AAU	GUG	CUG	GAG	UGG	CUG	GGA	GUU	UUA	UUA	UUA	UUA	UUA	UUA	UUA	UUA				
NQ7-3.3	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---				
-34.3	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---				
-1.3	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---				
-41.3	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---				
-5.3	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---				
-24.6	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---				
-51.3	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---				
-6.1	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---				
NQ10-2.2.5	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---				
-12.4.6	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---				
-15.9	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---				
-12.4.7	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---				

																			101	110	120	130
VH-Ox1	K	A	N	A	S	T	D	T	A	C	G	G	G	G	G	G	G	G				
K	AAA	AGG	AAU	AGU	CUU	CUU	GUU	GUU	GUU	GUU	GUU	GUU	GUU	GUU	GUU	GUU	GUU	GUU				
NQ7-3.3	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---				
-34.3	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---				
-1.3	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---				
-41.3	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---				
-5.3	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---				
-24.6	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---				
-51.3	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---				
-6.1	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---				
NQ10-2.2.5	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---				
-12.4.6	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---				
-15.9	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---				
-12.4.7	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---				

Fig. 1 Immunoglobulin heavy chain mRNA sequences (V regions) from anti-phenyl oxazolone hybridomas. The top line shows the sequence of the putative germ-line gene V_H -Ox1 joined to the D and J segments which were typically observed in the day 7 antibodies previously described³. The amino acid translation is given above, with numbering according to Kabat¹². The complementarity determining regions (CDR-1, -2, -3) are indicated, as are the D and J segments. The V_H region sequences of the individual hybridomas are given beneath. Dashes in the sequences indicate identity with the top line. An arrow marks probable identity. Upper and lower case letters are used to denote established and probable differences respectively. Where differences predict amino acid changes these changes are given above the relevant codon. A computer code¹⁹ is used to denote sequence ambiguities (5 = A or C, 6 = G or T, 7 = A or T, 8 = G or C). Two hybridomas (NQ7-51.3 and NQ7-6.1) express V region sequences which closely resemble the germ-line V_H gene expressed by the myeloma MC101²⁰.

Methods: Hybridomas against oxazolone were prepared as described previously. Fusions were made using the spleens from mice at 7 days (NQ2, NQ5 and NQ6) and 14 days (NQ7) after intraperitoneal (i.p.) primary immunization with 30 μ g i.p. alum precipitated antigen with 10⁹ killed *Bordetella pertussis* bacteria²². Secondary immunizations of antigen were given intravenously 3 days before fusion. All mice were of the BALB/c strain except mouse NQ6 (DBA/2). Sequencing strategy was essentially as described before³. Messenger RNA was prepared by a modification of the technique described by Favolora *et al.*²². 15–60 ng synthetic oligonucleotide primers were used to initiate cDNA synthesis on 1–2 μ g Poly(A)⁺ mRNA isolated from hybridomas. Reactions were carried out in the presence of 150 mM KCl, 50 mM Tris pH 8.3, 10 mM MgCl₂ with 1.5 μ M ³²P- or ³⁵S-labelled deoxynucleotide triphosphate and 50 μ M of the other three dNTPs. 2 units of reverse transcriptase, RT, (Anglian Biotechnology) was added per μ g of RNA and the reactions incubated at 42 °C for 15 mins. Reactions were 'chased' by addition of all four dNTPs (50 μ M, 10 mM Tris pH 8.3) and 1 unit RT per μ g RNA, and further incubation at 42 °C for 15 min. The reactions were terminated by addition to 30% formamide with dyes (0.1% xylene cyanol FF, 0.1% bromophenol blue, 10 mM EDTA) and boiled for 3–5 min prior to gel electrophoresis using 6% acrylamide gels.

seen are the only two which give rise to uncharged polar residues. In one instance a single variant sequence, differing from V_K -Ox1 at residues 34 and 36, is repeated in two hybridomas (NQ7-3.3 and NQ7-24.6). These same changes are also present in two other hybridomas (NQ10-12.4.6 and NQ11-7.12). However both of these show additional changes (Fig. 2). Five hybridomas (NQ7-1.3, NQ7-41.3, NQ10-2.2.5, NQ10-12.5 and NQ10-15.3) show identical nucleotide differences from the V_K -Ox1 germ-line sequence at positions 34 and 36. Four of these hybridomas show additional changes, some of which are repeated in yet other hybridomas. Hence the light chains show a pattern of highly related changes including repeats even in hybridomas originating from different animals and at different stages of the response.

In explaining this result, we cannot exclude the possibility that the repeats described at residues 34 and 36 originate from closely related germ-line genes which differ from V_K -Ox1 only at these two positions. But this could not explain all of the changes observed in the individual variants and one would have to conclude either that every variant was represented in the germ

line or that somatic mutation was also involved in producing the final unique pattern of changes seen in each variant light chain. For example, if we suppose that the light-chain sequence of NQ7-3.3 is present in the germ line then further changes are required to produce the sequences expressed by the hybridomas NQ10-2.2, -12.4 and NQ11-7.12. In addition this explanation would predict the presence of germ-line V genes differing from V_K -Ox1 by only two nucleotides. The study of BALB/c germ-line genes which are related to V_K -Ox1 permitted us to identify ten related sequences, one of which was identified as the predicted V_K -Ox1 sequence. None of the others were identical with the variant sequences described in this paper (J. Even and C. M., manuscript in preparation).

It seems more likely that most if not all of the changes described above arose by somatic point mutation of V_K -Ox1. The light chain changes are clustered around the first hypervariable region, a part of the antibody molecule which is likely to be directly involved in antigen contact⁷. In addition all of the repeated changes in this region predict amino acid changes in

OXAZOLONE KAPPA LIGHT CHAINS										10										20										30										40									
VK-Ox1										CAG										T										S										G									
NQ7-3.3										CAG										T										S										G									
-34.3										CAG										T										S										G									
-1.3										CAG										T										S										G									
-41.3										CAG										T										S										G									
-5.3										CAG										T										S										G									
-24.6										CAG										T										S										G									
-51.3										CAG										T										S										G									
-6.1										CAG										T										S										G									
NQ10-2.2.5										CAG										T										S										G									
-12.4.6										CAG										T										S										G									
-15.9										CAG										T										S										G									
-12.5										CAG										T										S										G									
NQ11-7.12										CAG										T										S										G									
-8.1										CAG										T										S										G									

Fig. 2 κ chain V region mRNA sequences from anti-phOx secreting hybridomas derived 14 days after primary immunization (NQ7-) and after secondary immunization (NQ10- and NQ11-). The top line shows the nucleotide sequence of the germline gene V_{κ} -Ox1. For other details see legend to Fig. 1. The primers used for mRNA sequencing have both been described. However, the 3' ends of some V regions were sequenced with alternative primers: C_{κ} -15_G (dTCACTCGTCAATTGT), which primes in the C_{κ} region between amino acids 122 and 126, and J_{κ} -15 (dGACCTCGATTGGCC), which primes between residues 104 and 108 of the rearranged J_{κ} 5 segment.

the immunoglobulin molecule and are therefore likely to affect the ability of the antibody molecule to bind antigen. Thus antigenic selection for variants with a better ability to bind antigen may account for the clustering of changes around the first hypervariable region. Some DNA stretches may also have a higher mutation rate than the average and behave as hot spots⁸, producing variants more frequently at certain positions. A high mutation rate (10^{-4} per cell per generation) in S107 myelomas at the fifth amino acid of the J_H segment has been reported⁹. The appearance of variants of V_H -Ox1 and V_{κ} -Ox1 during the maturation of the response to oxazolone is thus most likely the result of somatic mutation combined with antigenic selection.

This interpretation accounts for the correlation of the increase in affinity and appearance of somatic mutants with time after immunization. The primary immunizations described involved the use of adjuvant, resulting in a slow release of antigen and hence continual antigenic stimulation over a long time. Even the day 14 hybridomas are likely therefore to represent cells which have been subjected to multiple antigenic stimulations.

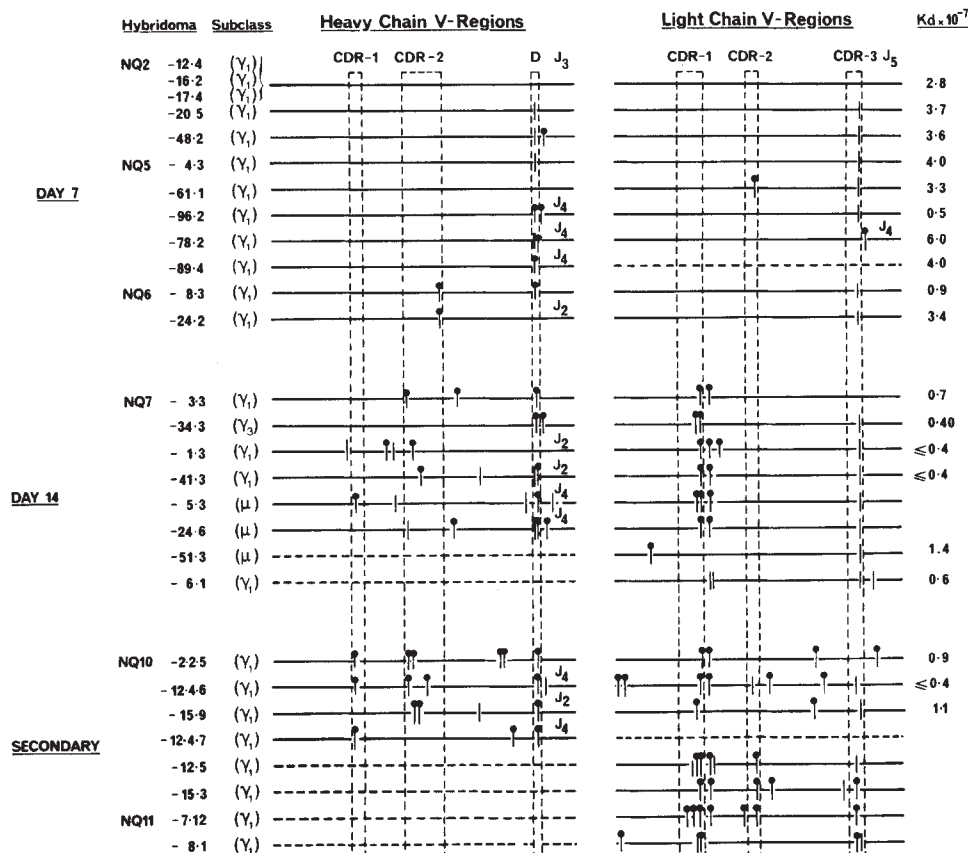
One unsolved problem is precisely when the variants arose. Two observations are relevant to this question. First, antibodies of the IgM class show clear examples of somatic mutation (for example, residue 104 of V_H in NQ7-5.3). Thus the IgM antibodies observed later in the response can express somatic mutations, whereas IgG antibodies observed early in the response are usually unmutated. Class switching is clearly not required

for mutations to occur. Second, variants can be observed at day 7, but at a much lower frequency. The single variant observed at day 7 shows only a single base change which does not improve the affinity for antigen. Time thus appears to be an important factor in the appearance of Ox1 variants with a higher affinity for oxazolone. Although it is possible that the variants already exist at day seven but are not localized within the spleen or are not part of the population that will fuse to form hybridomas, it is more straightforward to conclude that they arise after immunization and that their higher affinity for oxazolone eventually gives them a proliferative advantage. Our experiments do not address the question of whether antigenic stimulation actually increases the rate of mutation in B cells.

One intriguing observation concerning the affinity maturation of antibody responses is the importance of germ-line genes not found amongst the antibodies of the primary response. Indeed germ-line genes other than the V -Ox1 genes become dominant later in the response (Table 1). A full study of these sequences shows that they indeed belong to different families (C.B., G.M.G. and C.M., in preparation). Antisera produced later in the response to haptenic determinants show not only an increased affinity for antigen but also greater cross reactivity with other haptens¹¹. Whether this results from the appearance of variants or the use of the other genes remains to be elucidated.

The antibody responses to several other simple antigens have been investigated at the molecular level. These systems show

Fig. 3 Diagrammatic comparison of the mRNA sequences from anti-phOx secreting hybridomas derived at different stages after immunization with Ox-CSA. The hybridoma lines are the same as those described in Figs 1 and 2. The variable region sequences of each hybridoma has been compared with the sequences of V_H -Ox1 and V_K -Ox1 respectively. Unbroken horizontal lines denote sequence identity, broken lines represent extensive sequence differences. A vertical bar shows the relative position of codons containing nucleotide differences. A black circle indicates that these changes predict an amino acid difference at this position. Complementarity determining regions (CDR-1, -2, -3) have been marked as have the D and J regions. Where different J segments are observed these are represented according to the numbering of Sakano *et al.*⁴. Dissociation constants (K_d) were determined by fluorescent quenching as previously described²³. All values have now been calculated according to a computer program.



similar features to those described here¹²⁻¹⁶: a high degree of conservation in the overall V region structure, preferred combinations of V, D and J segments as well as of V_H and V_L ; and strong conservation of a particular amino acid at position 96 even though amongst all light chains this is the second most variable position of the κ chain¹². In addition putative somatic variants have been shown to have a higher affinity for antigen than their germ-line counterparts^{17,18,24}, although the published examples all show variations in the D segments which may not be somatic in origin. In this study the variant hybridoma, NQ7-1.3, expressed D and J segments which are typical of the low affinity antibodies observed at day seven, but itself produces an antibody with a high affinity for oxazolone (Fig. 3). Hence differences within the V segments must themselves account for the increase in affinity. The data presented here demonstrate that providing the basic structure of the antibody combining site is maintained, point mutations within the variable regions can produce antibodies with a better affinity for antigen. This increase in affinity is correlated with time, providing an explanation for the maturation effect.

We thank John Jarvis for technical assistance and John Sims for helpful discussions. G.M.G. was the recipient of an MRC studentship and an Usher Fellowship.

Received 5 July; accepted 6 September 1984.

- Siskind, G. W. & Benacerraf, B. *Adv. Immun.* **10**, 1-50 (1969).
- Tonegawa, S. *Nature* **302**, 575-581 (1983).
- Kaartinen, M., Griffiths, G. M., Markham, A. F. & Milstein, C. *Nature* **304**, 320-324 (1983).
- Max, E. E., Sidman, J. G. & Leder, P. *Proc. natn. Acad. Sci. U.S.A.* **76**, 3450-3454 (1979).
- Sakano, H., Maki, R., Kurosawa, Y., Roeder, W. & Tonegawa, S. *Nature* **286**, 676-683 (1980).
- Bothwell, A. L. M. *et al. Cell* **24**, 625-637 (1981).
- Amzel, L. M. & Poljak, R. J. *A. Rev. Biochem.* **48**, 961-997 (1979).
- Secher, D. S., Milstein, C. & Adetugbo, K. *Immun. Rev.* **36**, 51-72 (1977).
- Teillaud, J. L. *et al. Science* **222**, 721-726 (1983).
- Hooker, S. B. & Boyd, W. C. *Proc. Soc. exp. Biol. Med.* **47**, 187-190 (1941).
- Sperling, R., Francus, T. & Siskind, G. W. *J. Immun.* **131**, 882-885 (1983).
- Kabat, E. A., Wu, T. T., Bilofsky, H., Reid-Muller, M. & Perry, H. *Sequences of Proteins of Immunological Interest* (NIH, Bethesda, 1983).
- Rudikoff, S. *Contemp. Topics molec. Immun.* **9**, 169-209 (1983).
- Siegelman, M. & Capra, J. D. *Proc. natn. Acad. Sci. U.S.A.* **78**, 7679-7683 (1981).
- Slaughter, C. A. & Capra, J. D. *J. exp. Med.* **158**, 1615-1616 (1983).
- Rocca-Serra, J. *EMBO J.* **2**, 867-872 (1983).
- Reth, M., Hämmerling, G. J. & Rajewsky, K. *Eur. J. Immun.* **8**, 393-400 (1978).

- Rodwell, J. D., Gearhart, P. J. & Karush, F. J. *Immun.* **130**, 313-316 (1983).
- Staden, R. & McLachlan, A. D. *Nucleic Acids Res.* **10**, 141-156 (1982).
- Kataoka, T., Nikaido, T., Mijata, T., Moriaki, K. & Honjo, T. *J. biol. Chem.* **257**, 277-285 (1982).
- Kaartinen, M. *et al. J. Immun.* **130**, 937-945 (1983).
- Favaloro, J., Treisman, R. & Kamen, R. *Meth. Enzv.* **65**, 718-729 (1980).
- Parker, C. W. in *Experimental Immunology* 3rd edn (ed. Weir, D. M.) (Blackwell, London, 1978).
- Rothstein, T. L. & Gefter, M. L. *Molec. Immun.* **20**, 161-168 (1983).

Reactivity of HTLV-transformed human T-cell lines to MHC class II antigens

Nicole Suciu-Foca*, Pablo Rubinstein†, Mikulas Popovic‡, Robert C. Gallo‡ & Donald W. King§

* Columbia University, College of Physicians and Surgeons, 630 W. 168 St, New York, New York 10032, USA
 † The Lindsley F. Kimball Research Institute of the New York Blood Center, 310 E. 67 St, New York, New York 10021, USA
 ‡ National Institutes of Health, Building 37, 9000 Rockville Pike, Bethesda, Maryland 20205, USA
 § The University of Chicago, 950 E. 59 St, Chicago, Illinois 60637, USA

T-cell lines established from individuals infected with human T-cell leukaemia virus (HTLV) or generated by co-cultivation of normal human T cells with HTLV-infected T-cells, express class II (HLA-D/DR or Ia) antigens of the major histocompatibility complex (MHC)¹ and interleukin-2 (IL-2) receptors². Because the expression of these markers characterizes the differentiation of immunologically activated T cells³, we have now explored the possibility that HTLV-infected T cells might be primed to autologous or allogeneic Ia antigens expressed by the infecting cells. Our studies on the capacity of HTLV-infected T cells to display responses on mixed lymphocyte culture indicate that such

T cells as well as single-cell clones derived from them, react non-discriminatively to all known allelic variants of human HLA-D/DR antigens, including those expressed by the responding cells. This reaction is inhibited by antibody to human Ia and is not triggered by Ia-negative T-leukaemia cells. The structure recognized seems to be a common epitope determinant of human Ia antigens, as (HTLV-infected) T cells primed *in vitro* to one HLA-D/DR specificity display amplified responses to all other HLA-D/DR antigens. We therefore believe that autostimulation by a self-Ia determinant may trigger the clonal expansion of HTLV-infected T cells and potentiate autoimmune processes.

Figure 1 shows the magnitude of blastogenic responses in mixed lymphocyte culture (MLC) displayed by two HTLV-infected lines, FTH/TK (HLA-DR3, 6), derived by *in vitro* infection of human fetal thymocytes, and HUT-102 (HLA-DR2), derived from a patient with cutaneous T-cell lymphoma⁴. Both lines respond to HLA-D homozygous cells of any specificity and to B-lymphoblastoid cell lines generated by Epstein-Barr virus transformation of such HLA-D homozygous cells. Both FTH/TK and HUT-102 respond equally to self-like antigens on the HLA-D homozygous cells with which they share HLA-D/DR specificities. However, in MLC these cells do not respond to irradiated Ia-negative leukaemic T cell lines such as CEM and ERIC 3 (Fig. 1), suggesting that the stimulatory moieties are present exclusively on cells bearing MHC class II gene products.

In confirmation of this, we observed that these T-cell lines respond to B-lymphoblastoid cell lines and to alloactivated (Ia-positive) but not to resting (Ia-negative) T lymphocytes from the same individual, and that the MLC response is greatly

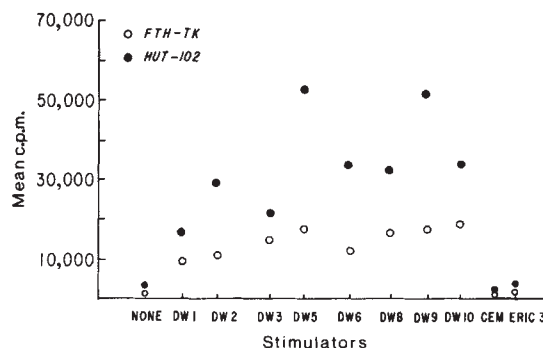


Fig. 1 MLC reactivity of FTH/TK and HUT-102 (1×10^3 cells per culture) to irradiated stimulating cells (5×10^4 cells per culture) from homologous T cells and from Ia-negative T cells, in 3-day MLC. The blastogenic response was quantitated by ^3H -thymidine incorporation and expressed as mean c.p.m. of triplicate reactions.

diminished if Ia antigens on the stimulating cells are blocked by preincubating the cells with the murine monoclonal antibody OKIa1 (Ortho Diagnostic), directed against human Ia determinants (Table 1). Monoclonal antibody Q1/18 to human MHC class I antigens has no inhibitory effect.

Because individual HLA-D/DR alleles are recognized by distinct T-cell clones^{5,6}, the universal anti-Ia reactivity of these cells indicates that the lines may be polyclonal or, if monoclonal, recognize a common structural determinant probably encoded by class II genes. To examine these alternatives, we have per-

Table 1 Effect of mouse monoclonal antibody to human Ia antigens (OKIa1) on the MLC responsiveness of HUT-102 and FTH/TK T-cell lines

Responding cells (1×10^3 cells per culture)	No. of stimulators	MLC responsiveness					
		Resting T cells		Alloactivated T lymphoblasts		β -Lymphoblastoid cell lines	
		a	b	a	b	a	b
HUT-102	1,234	1,650	1,420	15,460	6,820	18,320	5,210
FTH/TK	452	1,046	1,115	12,530	4,510	16,710	4,200

a, Untreated cells; b, antibody-treated cells. The effect of OKIa1 on responding or on stimulating cells (4,000 rad irradiated) was tested by preincubating 1×10^6 cells with the antibody ($10 \mu\text{l ml}^{-1}$) for 1 h at 37°C . Cultures were labelled with ^3H -thymidine and collected after 3 days. Results are expressed as mean c.p.m. of triplicate reactions. s.e.m. < 10%.

Table 2 Effect of BrdUrd and light illumination on MLC responsiveness of FTH/TK T cells

Primary stimulator	Secondary stimulators (irradiated HLA-D homozygous cells)						
	BrdUrd	None	DW3	DW9	DW8	DW10	DW1
None	No	862	6,998	6,646	6,760	9,292	5,798
None	Yes	502	5,272	5,120	5,708	7,200	3,054
DW3	No	3,498	19,394	25,402	27,002	21,976	19,422
DW3	Yes	660	6,706	5,996	6,694	5,550	5,234
DW9	No	1,818	16,834	17,972	14,968	15,780	17,494
DW9	Yes	756	5,348	4,252	4,586	3,462	2,918

FTH/TK cells were first sensitized by incubating 5×10^6 T cells with 5×10^6 irradiated DW3 or DW9 HLA-D homozygous cells. Cultures were freed of stimulating cells 24 h later by centrifugation on a 45% Percoll gradient and split into two parts. BrdUrd was added to half of each culture at a final concentration $5 \times 10^{-5}\text{M}$ and the incubation was continued for another 24 h. Cultures with and without BrdUrd were then exposed for 90 min to a Sylvania 40-W F40 D fluorescent light at a 15-cm distance. Dead cells were removed on a Ficoll-Hypaque gradient and viable cells were washed three times, counted and tested (at 1×10^3 cells per culture) for MLC reactivity to the original stimulator and to other HLA-D homozygous cells (5×10^4 cells per culture), in a 48-h assay. FTH/TK cells that had not been presensitized to any HLA-D homozygous cells were treated with BrdUrd and tested in identical conditions.

Table 3 MLC reactivity of T-cell subclones derived from FTH/TK to irradiated stimulating cells

Stimulating agent	Clone no.			
	7	8	15	21
None	760	758	548	918
DW1	4,916	5,508	4,724	4,712
DW2	4,582	7,198	5,484	7,382
DW3	7,398	7,468	7,382	9,200
DW4	6,898	5,648	4,088	5,580
DW5	6,270	7,402	9,318	6,596
DW8	9,428	7,524	7,972	6,590
DW9	14,392	11,508	9,428	14,672
DW10	12,020	12,232	10,760	10,936
FTH/TK	5,240	6,535	6,521	5,917
HUT-102	23,368	12,350	12,832	17,338

FTH/TK lymphocytes from the continuous culture were suspended at a concentration of 33 cells per ml of RPMI 1640 medium supplemented with 15% fetal calf serum and 20% IL-2. $100 \mu\text{l}$ of the cell suspension (0.33 cells per well) were dispensed in the wells of Falcon Microtest II trays on a feeder-layer consisting of irradiated human fibroblasts (MRC-5 line). Growing clones were selected and expanded in 24-well culture trays. Cells from different subclones were collected, washed and tested at 2×10^3 responding cells per culture (in medium without IL-2) for their reactivity to irradiated HLA-D homozygous T cells, HUT-102 and FTH/TK (5×10^9 stimulating cells per culture).

Table 4 Effect of monoclonal antibodies OKT3 and OKT4 on reactivity of HTLV-infected T cells

Responding T cells	Antibody added (0.25 µg per culture)	³ H-thymidine incorporation (mean c.p.m. of triplicate cultures)			
		Medium alone	Medium + 15% IL-2	Medium + irradiated stimulating cells (5 × 10 ⁴ cells per culture)	
FTH/TK (5 × 10 ³ cells per culture)	None	3,904	9,754	13,008	23,018
	OKT4	1,238	3,042	964	442
	OKT3	844	1,906	1,206	834
	OKT8	3,710	10,132	10,026	21,964
HUT/102 (5 × 10 ³ cells per culture)	None	13,996	23,372	82,146	104,530
	OKT4	214	2,074	312	320
	OKT3	424	1,140	130	252
	OKT8	11,534	28,856	74,418	98,342
Normal PBL (5 × 10 ⁴ cells per culture)	None	1,388	89,666	2,294	10,190
	OKT4	1,480	678	2,388	1,884
	OKT3	2,036	1,230	1,480	1,404

FTH/TK and HUT102 display the OKT3⁺ OKT4⁺ OKT8⁻ phenotype. Triplicate reactions containing responding cells from FTH/TK, HUT/102 or normal peripheral blood leukocytes (PBL) were tested for their reactivity to interleukin-2 (IL-2 containing phytaemagglutinin from Biological Products) or to irradiated stimulating cells consisting of either autologous cells or allogeneic PBL. Antibodies OKT3, OKT4 or OKT8 (Ortho) were added to the culture at the start of incubation. Cultures were labelled with tritiated thymidine after 72 h and collected 16 h later.

formed 5-bromodeoxyuridine (BrdUrd) 'suicidal' experiments which are based on the capacity of BrdUrd to become incorporated into the DNA of dividing lymphocytes, rendering them light sensitive⁵ (Table 2). The MLC reactivity of the 'non-primed' cells that had survived following exposure to BrdUrd and light is similar to that of cells not pretreated with the drug. Reactivity is not clonally restricted, as FTH/TK cells primed for 24 h to DW3 or DW9 HLA-D homozygous cells displayed a significantly stronger response, reminiscent of a memory reaction, not only when re-exposed to the original stimulator, but also when challenged with other HLA-D homozygous cells. Following BrdUrd treatment, the MLC responsiveness of the surviving cells returned to the pre-sensitization level. Furthermore the finding that both cell lines, that is, cells carrying the T-cells' own HLA-D/DR antigen (DW3) and cells of a different specificity (DW9), induce the FTH/TK cells to display amplified, memory-type responses in secondary MLC and that such responses are abolished by BrdUrd treatment, suggests that a common class II structure may be responsible. Single-cell clones derived from FTH/TK maintain the universal anti-HLA-D/DR alloreactivity together with the capacity to react to autologous stimulating cells (Table 3). Similar to normal T lymphocytes, HTLV-infected T-cell lines show greatly diminished responses to IL-2 and to self or allogeneic Ia antigens when tested in the presence of monoclonal antibodies OKT3 and OKT4, which detect surface molecules involved in T-cell reactivity to MHC class II antigens⁷. Monoclonal anti-OKT8 antibodies, used as a control, have no effect on proliferation. The rate of blastogenesis displayed by FTH/TK, and particularly by HUT-102 in the absence of any stimulating cells, is also inhibited by OKT3 and OKT4 antibodies. This may be due to suppression of the growth-promoting anti-self-Ia responsiveness of these cells.

Our findings suggest that normal MHC class II molecules, including 'self', trigger the blastogenesis of HTLV-infected T cells. In turn, this indicates that immunological recognition of 'self' may be altered by the virus and suggests explanations for the paradoxical hypergammaglobulinaemia of patients suffering from acquired immune deficiency syndrome (AIDS)—characteristically deficient in T-helper cells. Autostimulation by self-Ia may cause direct potentiation of autoimmune processes as well as expansion of HTLV-infected malignant cells.

Received 9 May; accepted 2 September 1984.

1. Gallo, R. C. *et al. Proc. natn. Acad. Sci. U.S.A.* **79**, 5680–5683 (1982).
2. Popovic, M. *et al. Proc. natn. Acad. Sci. U.S.A.* **80**, 5402–5406 (1983).
3. Suciu-Foca, N. *et al. Transplantation* **10**, 845–848 (1978).
4. Poiesz, B. J. *et al. Proc. natn. Acad. Sci. U.S.A.* **77**, 7415–7419 (1980).
5. Zoschke, D. & Bach, F. H. *J. immun. Meth.* **1**, 55–65 (1971).
6. Suciu-Foca, N. *et al. J. exp. Med.* **156**, 283–287 (1982).
7. Reinherz, E. L. *et al. Immun. Rev.* **74**, 83–112 (1983).

Isolation of a cDNA clone encoding rat insulin-like growth factor-II precursor

Harvey J. Whitfield*, Carmelo B. Bruni†, Rodolfo Frunzio†, Jeffrey E. Terrell*, S. Peter Nissley‡ & Matthew M. Rechler*

* Laboratory of Biochemical Pharmacology, National Institute of Arthritis, Diabetes and Digestive & Kidney Diseases, National Institutes of Health, Bethesda, Maryland 20205, USA

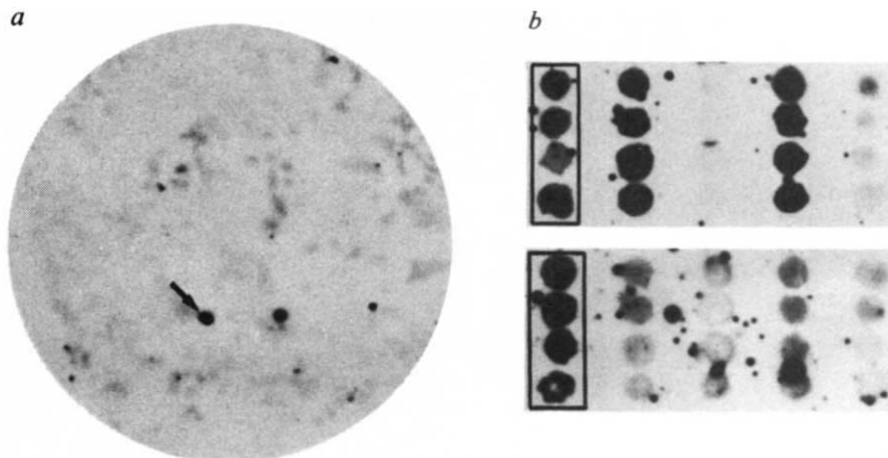
† Centro di Endocrinologia ed Oncologia Sperimentale del CNR, Istituto di Patologia Generale, II^a Facoltà di Medicina e Chirurgia, Naples, Italy

‡ Metabolism Branch, National Cancer Institute, National Institutes of Health, Bethesda, Maryland 20205, USA

Insulin-like growth factor-I (IGF-I) and IGF-II are mitogenic polypeptides of relative molecular mass (M_r) ~7,500 isolated from human plasma^{1,2} each containing four peptide domains in a single chain and identical at more than 60% of their amino acid loci. The B- and A-domains of the IGFs are ~40% identical to the B- and A-chains of human insulin^{1,2}. IGF-I and IGF-II have similar *in vitro* biological activities² and receptor reactivity³, but are immunologically distinct^{4,5}. IGF-I appears to mediate the effects of growth hormone on cartilage to promote skeletal growth^{5,6}, whereas IGF-II may have a special role in fetal development^{7,8} and in the central nervous system⁹. To investigate the *in vivo* role of IGF-II, we have studied IGF-II biosynthesis in the BRL-3A rat liver cell line¹⁰. BRL-3A cells synthesize and secrete a 7,484 M_r protein 93% identical to human IGF-II and representing rat IGF-II (rIGF-II)¹¹. Rat IGF-II is synthesized as a ~22,000 M_r prepro-rIGF-II (ref. 12) from 12S poly(A)⁺mRNA¹³. In addition, ~20,000 M_r pro-rIGF-II has been identified in lysates of biosynthetically labelled intact BRL-3A cells¹⁴. We report here the isolation of an almost complete cDNA clone for rIGF-II. Our results indicate that pro-rIGF-II is synthesized as a 156 amino acid peptide precursor (17,619 M_r) containing mature rIGF-II 1–67 at its amino-terminus and an 89-residue carboxy-terminal peptide extension.

A cDNA library was constructed from 12S poly(A)⁺ mRNA from BRL-3A cells (see Fig. 1 legend) and screened using a 5' [γ -³²P]-labelled tetradecamer (designated 14-mer A) corresponding to amino acids 44–48 of rat IGF-II¹¹. Twenty out of ~10,000 transformants screened at high cell density gave a strong autoradiographic signal (Fig. 1a). Positive transformants were

Fig. 1 Identification of a rat IGF-II cDNA clone. *a*, One filter from which the prIGF-II-1 clone (arrow) was identified. *b*, Autoradiograms of clones replicated on nitrocellulose filters screened with 5'-labelled 14-mer A (top) or 14-mer B2 (bottom). The boxed areas show quadruplicate copies of single isolates of the clone containing prIGF-II-1. Hybridization to four other clones that were positive with probe 14-mer A after low temperature washes is also shown (four isolates of each are plated in vertical rows). After washing at 38 °C, two of four clones were positive with probe 14-mer A, but none of the four was positive with probe 14-mer B2.



Methods: When poly(A)-containing RNA from BRL-3A cells was fractionated on a 10–30% sucrose gradient, RNA directing the translation of ~22,000M_r prepro-rIGF-II was most abundant in the 12S fraction^{13,28}. This 12S RNA was used to establish cDNA clones by the procedure of Land *et al.*¹⁸ with modifications. Single-stranded cDNA was synthesized using oligo(dT)_{12–18} (P-L Biochemicals) as primer, 50 µg ml⁻¹ actinomycin D and avian myeloblastosis virus (AMV) reverse transcriptase (Beard, Life Sciences)¹⁹. The single-strand cDNA was then tailed with poly(dC) using terminal transferase (P-L Biochemicals)²⁰, annealed with oligo(dG)_{12–18} primer (P-L Biochemicals) and the complementary DNA strand synthesized using AMV reverse transcriptase¹⁹. The double-stranded cDNA was tailed with poly(dC), annealed with dG-tailed *Pst*I-cut pBR322 plasmid (NEN), and used to transform *Escherichia coli* strain MC 1061 (ref. 21) made competent by the CaCl₂ procedure²². From 3 µg of size-selected poly(A)⁺ mRNA, 700 ng of double-stranded cDNA were obtained, of which 360 ng were used to obtain ~10,000 tetracycline-resistant clones. These clones were screened on a series of duplicate 137-mm nitrocellulose filters (HATF, Millipore) according to the high density screening procedure of Hanahan and Meselson²³. Three mixtures of tetradecamer oligonucleotide probes (synthesized by P-L Biochemicals) complementary to two different portions of the 67 amino acid rIGF-II sequence¹¹ were used to screen the filters. The first set, designated 14-mer A, 5'-d(AA₂CA₂CA₂TC₂TC)-3', corresponds to amino acid residues 44–48 (Glu-Glu-Cys-Cys-Phe, excluding the third nucleotide of the Phe codon) and contains equimolar amounts of the 16 sequences. The second set corresponds to residues 57–61 (Glu-Thr-Tyr-Cys-Ala, excluding the third nucleotide of the Ala codon) and was synthesized as two pools of oligonucleotides, 5'-d(GC₂CA₂TA₂GT₂TC)-3' and 5'-d(GC₂CA₂TA₂GT₂TC)-3', each containing 16 equimolar sequences and designated 14-mer B1 and 14-mer B2, respectively. The 14-mer A probe was 5'-labelled using T4 polynucleotide kinase and [γ -³²P]ATP¹⁹, and used to screen the cDNA library. Conditions for hybridization and washing of filters were as described for screening with mixtures of short oligonucleotide probes²³. The filters were washed four times at 25 °C for 5 min and then at 30, 32 and 34 °C for 10 min each in 6 × SET (1 × SET is 0.15 M NaCl, 0.001 M EDTA, 0.03 M Tris-HCl, pH 8.0) before autoradiography. Potentially positive clones were identified on duplicate filters. Areas containing possible positive clones were picked, plated at a low cell density and rescreened with 5'-labelled 14-mer A probe. Positive clones were picked, purified, plated in an ordered array, replicated on nitrocellulose filters, and screened with 14-mer A, B1 and B2 probes. In addition to the washes described above, filters were washed at higher temperatures (2–3 °C intervals up to 47 °C, which exceeds the theoretical temperature of denaturation for hybrids of the 14-mer A, B1 and B2 probes²⁴). Clone prIGF-II-1 gave a positive hybridization signal with the 14-mer A and 14-mer B2 probes after washes up to 43 and 47 °C for 10 min, respectively. The 14-mer B1 probe hybridized to prIGF-II-1 after washes at lower temperatures, but the hybridization signal decreased with increasing wash temperature, disappearing at 44 °C (data not shown).

purified and rescreened with labelled tetradecamers B1 and B2 corresponding to amino acids 57–61 of rat IGF-II. One clone, designated prIGF-II-1, gave a positive hybridization signal with both probe 14-mer A (Fig. 1*b*, top) and probe 14-mer B2 (Fig. 1*b*, bottom) at high stringency washes; probe 14-mer B1 was positive only in less stringent wash conditions (data not shown). Plasmid DNA prepared from prIGF-II-1 was digested with *Pst*I, fractionated on agarose gels, transferred to nitrocellulose filters and the filters hybridized with 5'-labelled 14-mer A and 14-mer B2. A DNA fragment of 810 base pairs (bp) corresponding to the cDNA inserted in the *Pst*I site was observed. This insert hybridized with both the 14-mer A and 14-mer B2 probes (data not shown).

The prIGF-II-1 plasmid DNA was digested with restriction enzymes *Pvu*II, *Eco*RI or *Bam*HI and electrophoresed on agarose gels. The DNA fragments were blotted onto nitrocellulose and hybridized with the 14-mer A or 14-mer B2 probes (Fig. 2*a*). Both probes hybridized with the same *Eco*RI (5.2 kilobases, kb) and *Bam*HI (3.8 kb) restriction fragments. In contrast, the two probes hybridized with different *Pvu*II fragments: the 14-mer A probe hybridized with the 1.6 kb fragment, whereas the 14-mer B2 probe hybridized with the 3.5 kb fragment. Hybridization of 14-mer A and 14-mer B2 probes to the short and long *Pvu*II fragments, respectively, is readily explained by the presence of a *Pvu*II site in pBR322 (nucleotide 2,067) and of a potential *Pvu*II site in the cDNA corresponding to serine 50 (Fig. 2*b*). A detailed restriction map of this plasmid was derived in order to identify overlapping DNA fragments suitable for 5'-labelling and nucleotide sequence analysis (Fig. 2*b*).

The complete nucleotide sequence of 752 bases (minus dG-

and dC-tails) is presented in Fig. 3*a*. From the 5'-end, long open reading frames exist in all three frames: 357 nucleotides (Fig. 3*a*), 288 nucleotides (+1 frame) and 372 nucleotides (–1 frame). The translation product deduced from the first 89 nucleotides in the frame shown, however, corresponds to amino acids 38–67 of rIGF-II (ref. 11), representing part of the C-domain (residues 38–40), all of the A-domain (residues 41–61) and all of the D-domain (residues 62–67). Glu 67 is followed by 89 additional amino acid residues before encountering a TGA stop codon after nucleotide 357. This 89 residue peptide extension includes three pairs of dibasic amino acids and 12 single lysines or arginines, possibly accounting for the multiplicity of rIGF-II species of different size and charge^{10,11,14}. Unlike rIGF-II 1–67, the peptide extension does not contain cysteine. The single TGA stop codon at the end of the pro-rIGF-II precursor (positions 358–360) is followed by an additional 392 nucleotides. A canonical polyadenylation recognition sequence, AATAAA¹⁵, has not been identified; however, a possible poly(A) addition site (ATTAAG) is present at positions 707–712).

Primer extension experiments were performed to determine the length of the 5'-region of the mRNA that is missing from our clone and to obtain further structural information. A 5'-labelled *Dde*I–*Pvu*II fragment (72 bases long, Fig. 2*b*) was hybridized to poly(A)⁺ RNA from BRL-3A cells and the DNA strand of the hybrid extended with reverse transcriptase. Analysis of the products on polyacrylamide gels showed the presence of a single labelled band ~380 nucleotides long (ref. 28). The extended DNA fragment was eluted from the gel and sequenced. The DNA sequence obtained corresponding to amino acids 3–40 of rIGF-II is shown in Fig. 3*b* and represents the complete sequences of the C-domain (residues 33–40) and

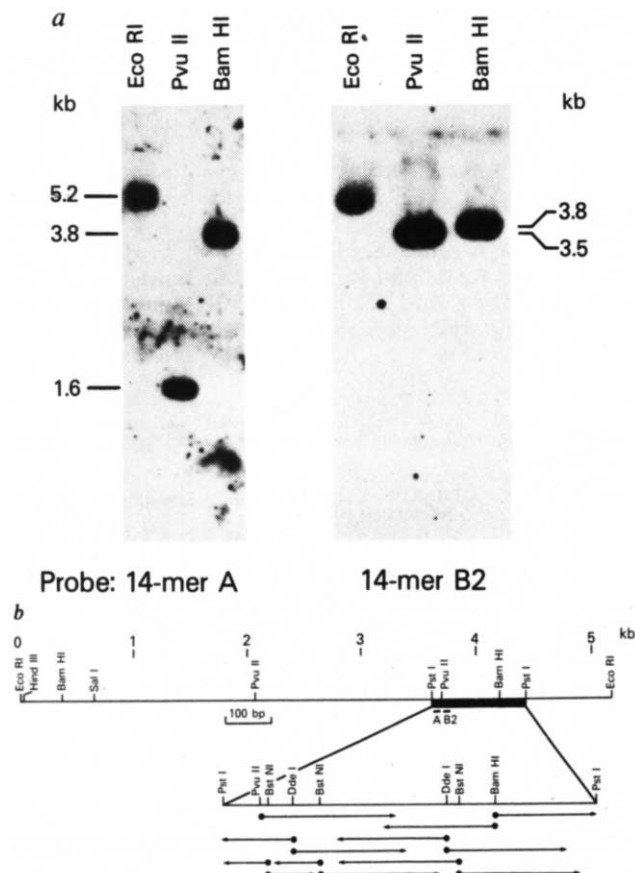


Fig. 2 Structural analysis of the prIGF-II-1 clone. *a*, Southern blots of prIGF-II-1 plasmid DNA digested with the indicated restriction enzymes and hybridized with the 5'-[γ - 32 P]-labelled 14-mer A (left) or 14-mer B2 (right) oligonucleotide probes. *b*, Restriction map of the prIGF-II-1 plasmid DNA linearized at the *Eco*RI site. Bold line, cDNA insert. The positions of the *Pvu*II and *Bam*HI sites are indicated, as well as the relative positions of the 14-mer A and 14-mer B2 oligonucleotides. The restriction map of the *Pst*I cDNA insert is shown on an expanded scale to indicate the overlapping DNA fragments used for sequence determination. Arrows, the size of each fragment and direction in which it was sequenced. Closed circles, position of the 32 P label at the 5'-end of the fragment.

Methods: Plasmid DNA was purified²⁵, digested with the indicated restriction enzymes, subjected to electrophoresis on three 1% agarose gels, and blotted onto nitrocellulose by standard methods¹⁹. *Hind*III-digested λ DNA and *Hae*III-digested ϕ X174 DNA were used as markers to determine the sizes of the restriction fragments. The blots were hybridized with the 5'-[γ - 32 P]-labelled 14-mer A, 14-mer B1 (not shown), or 14-mer B2 oligonucleotide probes. Hybridizations were at 35 °C in 6 $\times$ NET (1 $\times$ NET is 0.15 M NaCl, 0.001 M EDTA, 0.015 M Tris-HCl, pH 7.5), 5 $\times$ Denhardt's solution (1 $\times$ Denhardt's solution is 0.2 mg ml⁻¹ Ficoll, 0.2 mg ml⁻¹ polyvinylpyrrolidone, 0.2 mg ml⁻¹ bovine serum albumin), 0.5% Nonidet P-40, and 50 μ g ml⁻¹ yeast tRNA (BRL). Filters were washed in 6 $\times$ SSC (1 $\times$ SSC is 0.15 M NaCl, 0.015 M sodium citrate, pH 7.2) up to 39 °C as described in Fig. 1 legend.

of most of the B-domain (residues 3–32). The amino acid sequence deduced from this nucleotide sequence confirms the protein sequence of mature rIGF-II reported by Marquardt *et al.*¹¹.

Radiosequence analysis of ~20,000 *M_r* pro-rIGF-II purified from biosynthetically labelled BRL-3A cells, has established that mature rIGF-II 1–67 is located at the amino terminus of ~20,000 *M_r* pro-rIGF-II (ref. 16). Together with the nucleotide sequences of our clone and the primer-extended DNA fragment, these results indicate that pro-rIGF-II is a protein of 156 amino acids (17,619 *M_r*). These results also establish the following structure for rIGF-II mRNA: a ~100 nucleotide 5'-untranslated region, a 72-base segment encoding a prepeptide, 468 bases

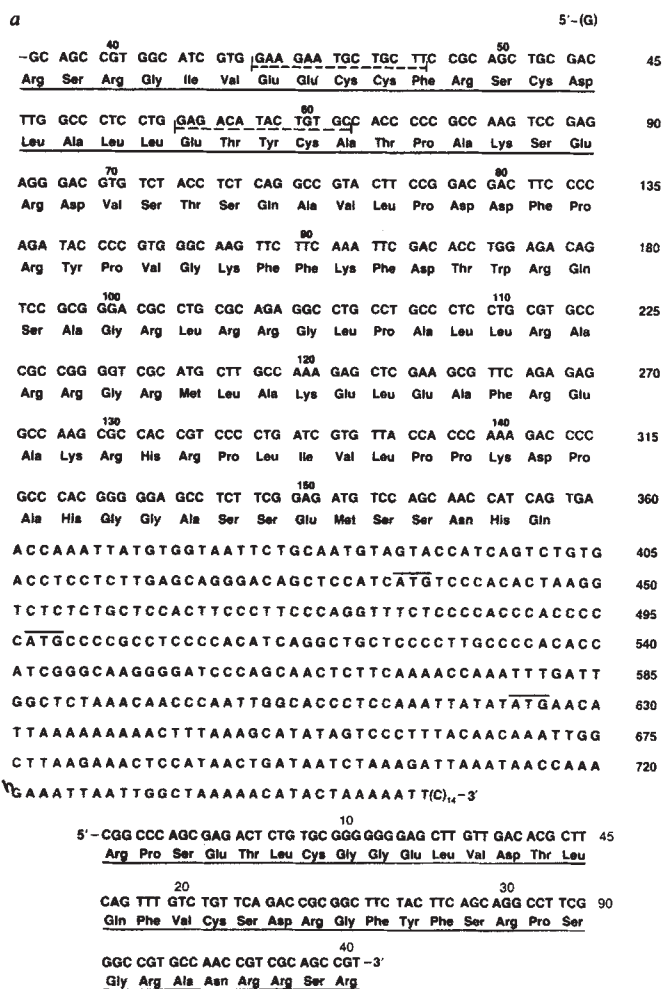


Fig. 3 *a*, Nucleotide sequence of the prIGF-II-1 clone showing the open reading frame from nucleotide 1–357 and the deduced amino acid sequence in this frame. The region corresponding to amino acids 38–67 of mature rat IGF-II, *M_r* 7,484 (ref. 11), is underlined (solid line). Nucleotide residues are numbered in the 5' to 3' direction starting with the first trinucleotide codon appearing in the prIGF-II-1 clone. Numbers to the right refer to nucleotide numbers. The identity of the 5'-base of the first nucleotide in the prIGF-II-1 sequence following the 5'-dG tail was ambiguous and thus is represented by a dash. The numbers above the nucleotide sequence are referenced to the amino acid sequence of 7,484 *M_r* rIGF-II (ref. 11). The overlined ATG codons indicate methionine codons that might initiate three potential peptides encoded by the 3'-region of the prIGF-II-1 clone. The underlined nucleotides (dashed line) designate the sequences homologous to probes 14-mer A and B2, respectively. *b*, Nucleotide sequence of the primer-extended cDNA. The sequence of the DNA region corresponding to amino acids 3–40 of mature rat IGF-II is shown.

Methods: *a*, Plasmid DNA was cleaved with either *Bam*HI or *Pvu*II, labelled at the 5'-ends with T4 polynucleotide kinase and [γ - 32 P]ATP¹⁹ and redigested with *Pst*I. The 5'-labelled *Pvu*II fragment (735 bp) and *Bam*HI fragments (595 and 220 bp) were isolated on 6% polyacrylamide gels, recovered by electroelution¹⁹ and sequenced according to Maxam and Gilbert²⁶. Plasmid DNA also was digested with *Pst*I and the cDNA insert isolated on 5% polyacrylamide gels. The purified *Pst*I fragment was cleaved with either *Dde*I or *Bst*NI and labelled at the 5'-ends. The strands were separated¹⁹ and both strands sequenced. *b*, The purified 810 bp *Pst*I fragment from prIGF-II-1 was cleaved with *Dde*I, terminally labelled at the 5'-end¹⁹ and digested with *Pvu*II. A 72-base *Dde*I-*Pvu*II fragment (nucleotides 39–110, inclusive) was isolated on a 10% denaturing polyacrylamide gel²⁶ and hybridized to poly(A)⁺ RNA isolated from BRL-3A cells. DNA-RNA hybrids were extended with reverse transcriptase²⁷ and sized on a 6% denaturing polyacrylamide gel. The 380-base extended DNA fragment was eluted and purified. We determined the sequence of the first 215 bases of the primer-extended fragment. Only those nucleotides corresponding to amino acids 3–40 of rIGF-II are shown.

encoding pro-rIGF-II, and a 392-base 3'-untranslated region. The size determined for rIGF-II mRNA (~1,303 bases) agrees well with that predicted for the functional 12S mRNA used to generate our cDNA library.

A computer search of the Los Alamos National Laboratory Genbank for nucleotide sequences potentially homologous to rIGF-II cDNA revealed that the most marked homologies were to the A- and B-domains of human IGF-II²⁹ and IGF-I¹⁷, with weaker homologies to the B- and A-domains of insulins from several species.

Availability of the prIGF-II-1 clone, corresponding to functional rIGF-II mRNA, should help elucidate the regulation of IGF-II gene expression and its possible role in development^{7,8} and in nervous system function⁹.

Received 6 June; accepted 27 September 1984.

1. Rinderknecht, E. & Humbel, R. E. *J. biol. Chem.* **253**, 2769-2776 (1978).
2. Zapf, J., Froesch, E. R. & Humbel, R. E. *Curr. Topics cell. Regul.* **19**, 257-309 (1981).
3. Rechler, M. M. et al. in *Insulin-like Growth Factors/Somatomedins: Basic Chemistry, Biology and Clinical Importance* (ed. Spencer, E. M.) 459-490 (de Gruyter, New York, 1983).
4. Van Wyk, J. J., Svoboda, M. E. & Underwood, L. E. *J. clin. Endocr. Metab.* **50**, 206-208 (1980).
5. Zapf, J., Walter, H. & Froesch, E. R. *J. clin. Invest.* **68**, 1321-1330 (1981).
6. Schoenle, E., Zapf, J., Humbel, R. E. & Froesch, E. R. *Nature* **296**, 252-253 (1982).
7. Moses, A. C. et al. *Proc. natn. Acad. Sci. U.S.A.* **77**, 3649-3653 (1980).
8. Adams, S. O., Nissley, S. P., Handwerker, S. & Rechler, M. M. *Nature* **302**, 150-153 (1983).
9. Haselbacher, G. K., Schwab, M. E., Pasi, A. & Humbel, R. E. *Symposium Insulin-like Growth Factors/Somatomedins*, Nairobi, Kenya, 13-15 November, p. 75 (abstract) (1982).
10. Nissley, S. P. et al. in *Insulin-like Growth Factors/Somatomedins: Basic Chemistry, Biology, and Clinical Importance* (ed. Spencer, E. M.) 31-48 (de Gruyter, New York, 1983).
11. Marquardt, H., Todaro, G. J., Henderson, L. E. & Oroszlan, S. *J. biol. Chem.* **256**, 6859-6865 (1981).
12. Acquaviva, A. M., Bruni, C. B., Nissley, S. P. & Rechler, M. M. *Diabetes* **31**, 656-658 (1982).
13. Yang, Y. W.-H. et al. in *Insulin-like Growth Factor/Somatomedins: Basic Chemistry, Biology, and Clinical Importance* (ed. Spencer, E. M.) 603-610 (de Gruyter, New York, 1983).

Whilst this manuscript was being reviewed, Dull et al.³⁰ reported the nucleotide sequence of cDNA clones established from total poly(A)⁺ BRL-3A RNA encoding pre-pro-rIGF-II. Their clones are colinear with ours over ~600 bases and have virtually identical nucleotide sequences, but, unlike ours, they contain a 5'-untranslated region of 1,064 nucleotides. The deduced length of their RNA, ~2,000 bases, is approximately twice as long as predicted for a functional 12S rIGF-II mRNA^{13,28}.

We thank D. E. Graham, M. Kanehisa, Y. W.-H. Yang and J. A. Romanus for helpful discussions and assistance. H.J.W. was supported by a Ford Foundation Senior Minority Postdoctoral Fellowship while on sabbatical leave from the University of Michigan.

14. Yang, Y. W.-H., Romanus, J. A., Liu, T.-Y., Nissley, S. P. & Rechler, M. M. *J. biol. Chem.* (in the press).
15. Berget, S. M. *Nature* **309**, 179-182 (1984).
16. Yang, Y. W.-H., Rechler, M. M., Nissley, S. P. & Coligan, J. E. *J. biol. Chem.* (in the press).
17. Jansen, M. et al. *Nature* **306**, 609-611 (1983).
18. Land, H., Grez, M., Hauser, H., Lindenmaier, W. & Schutz, G. *Meth. Enzym.* **100**, 285-292 (1983).
19. Maniatis, T., Fritsch, E. F. & Sambrook, J. *Molecular Cloning* (Cold Spring Harbor Laboratory, New York, 1982).
20. Michelson, A. M. & Orkin, S. H. *J. biol. Chem.* **257**, 14773-14782 (1982).
21. Casadaban, M. J. & Cohen, S. N. *J. molec. Biol.* **138**, 179-207 (1980).
22. Mandel, M. & Higa, A. *J. molec. Biol.* **53**, 159-162 (1970).
23. Hanahan, D. & Meselson, M. *Meth. Enzym.* **100**, 333-342 (1983).
24. Suggs, S. V. et al. in *Developmental Biology Using Purified Genes* (eds Brown, D. D. & Fox, C. F.) 683-693 (Academic, New York, 1981).
25. Clewell, D. B. & Helinski, D. R. *Biochemistry* **9**, 4428-4440 (1970).
26. Maxam, A. M. & Gilbert, W. *Meth. Enzym.* **65**, 499-560 (1980).
27. Wickens, M. P., Buell, G. N. & Schimke, R. T. *J. biol. Chem.* **253**, 2483-2495 (1978).
28. Rechler, M. M. et al. *Cancer Cells* **3** (in the press).
29. Bell, G. I. et al. *Nature* **310**, 775-777 (1984).
30. Dull, T. J., Gray, A., Hayflick, J. S. & Ullrich, A. *Nature* **310**, 777-781 (1984).

Regulation of heat shock protein 70 gene expression by c-myc

Robert E. Kingston, Albert S. Baldwin Jr & Phillip A. Sharp

Center for Cancer Research and Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA

The *myc* gene seems to have a causal role in tumour formation in man, mouse and avian systems¹⁻³. The *myc* gene product has been localized to the nucleus^{4,5}, suggesting that it may be involved in the regulation of gene expression. The level of expression of the mammalian heat shock protein 70 (HSP70) gene is elevated in several tumour cell lines⁶, implying that a cellular function expressed in these tumour lines can stimulate HSP70 production. We report here that the gene product of a rearranged mouse c-myc gene is capable of stimulating expression of chimaeric genes containing a *Drosophila hsp70* promoter region and 5'-flanking sequences. This stimulation is dependent on sequences located more than 200 bases 5' of the normal start of *hsp70* transcription.

To investigate *myc* activity we used a construction in which the second and third exons of a *myc* gene isolated from a mouse plasmacytoma² are expressed from the simian virus 40 (SV40) early promoter (termed pSV2-cmyc, Fig. 1a and ref. 7).

To assess the effect of the *myc* gene on *hsp70* expression, we used hybrids containing the promoter regions of two of the five genes in *Drosophila* which encode HSP70 (refs 8-10); these are linked to coding sequences for either dihydrofolate reductase (DHFR) or chloramphenicol acetyltransferase (CAT)^{11,12}. One of the promoter regions we have investigated is contained at the proximal end of the *Drosophila* 87Cl loci⁸ (plasmids prefixed pPHS, Fig. 1b), the other is located at the distal end of the same loci (plasmid pDHSP-cat, Fig. 1b and ref. 13).

Using a stable transfection protocol¹⁴, the effect of *myc* on expression of the *dhfr* hybrids was tested by transfecting a *dhfr*-deficient CHO cell line^{15,16} with the *hsp70dhfr* test vector

in the presence or absence of pSV2-cmyc. Two days later, cells were placed in selective media (lacking exogenous thymidine, adenine or deoxyadenine) at one of two levels; the more stringent of the two contained 0.005 μ M methotrexate, an inhibitor of *dhfr* function. The number of colonies obtained in this protocol after transfection of the *dhfr* test vector alone reflects the basal level of expression of the chimaeric gene. As expected of a gene that stimulates chimaeric gene expression, co-transfection of pSV2-cmyc with pPHSP-dhfr results in an increase in the number of colonies obtained (Fig. 2a), suggesting that the *myc* gene product stimulates expression from pPHSP-dhfr.

In a further experiment, sequences were deleted from the 5' side of the promoter region of pPHSP-dhfr to either the *Hind*III site at -780 bases (to create pPHSH3-dhfr) or to the *Xho*I site at -200 bases (pPHSX1-dhfr). Co-transfection of pSV2-cmyc with pPHSH3-dhfr resulted in an increase in the number of colonies obtained in the stable co-transfection protocol (Fig. 2b), while co-transfection of pSV2-cmyc with pPHSX1-dhfr had no such effect (Fig. 2c). These results suggest that sequences contained between -200 and -780 bases are necessary for regulation of pPHSP-dhfr expression by the *myc* protein.

When DNA is transfected into cells, it can form long aggregates that become integrated into chromosomal DNA. This property raises the possibility that the stimulation we have observed is caused by the pSV2-cmyc sequences acting in *cis*, and not by the *myc* gene product. To address this question, a 4-base deletion was introduced at one of the two *Pst*I sites in the second *myc* exon to produce psmyc-20 (see Fig. 1a). This deletion will produce a frameshift in the *myc* coding region and the vector will thus make a truncated *myc* protein product while maintaining virtually the entire pSV2-cmyc DNA sequence. This mutant does not stimulate the expression of either pPHSH3-dhfr or pPHSP-dhfr (Fig. 2d, e), implying that the observed stimulation is caused by the *myc* gene product.

Although co-transfection with pSV2-cmyc consistently stimulated colony formation, the level of stimulation varied from two- to eightfold and the amount of pSV2-cmyc necessary to achieve stimulation varied between 0.4 and 2.0 μ g DNA. A potential

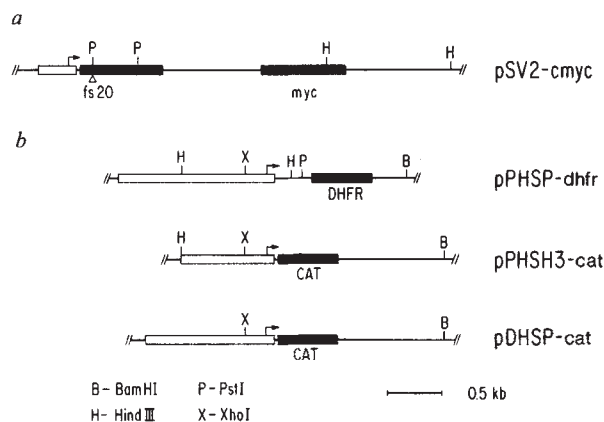


Fig. 1 Schematic representations of vectors used. Open blocks designate promoter regions, filled blocks indicate the coding sequences of the expressed gene; arrows indicate the normal transcriptional start site for each promoter region. Plasmid pSV2-cmyc (a) contains the SV40 early promoter fused to the second and third exons of a c-myc gene isolated from a mouse plasmacytoma^{2,7}. Plasmids pHSP-dhfr and pHSH3-cat (b) both contain the promoter region from the proximal *hsp70* promoter of *Drosophila* 87C1 (from plasmid pPW232)⁸. Plasmid pDHSP-cat (pDHSP-cat 1 of ref. 13) contains the proximal *hsp70* promoter of the two distal *hsp70* genes at *Drosophila* 87C1. Both *dhfr* and *cat* vectors contain a 5' and 3' splice site and the SV40 early transcription unit polyadenylation sequence.

Methods: Plasmid constructions. pfsmyc-20: Plasmid pSV2-cmyc was partially digested with *Pst*I and linearized plasmid was isolated, then incubated with T4 DNA polymerase and dATP, dGTP, dCTP and TTP, producing a blunt end at the *Pst*I site; the plasmid was circularized with T4 DNA ligase. These manipulations resulted in a 4-base deletion at the indicated sequence as determined by sequencing of the DNA²⁸. pHSP-dhfr: The *hsp70* promoter-containing *Pst*I fragment (extending from +85 to 1.3 kb upstream of the initiation site) of plasmid p232.3 (ref. 8) was isolated and blunt ends were made. This fragment was cloned into *Pvu*II-cleaved pAdD26SV(A) (ref. 11). For historical reasons, this vector thus contains the adenovirus major late promoter and SV40 origin upstream of the *hsp70* promoter. These sequences were deleted from all subsequent vectors. pHSH3-dhfr: The 5.4-kb band of a partial *Hind*III digest of pHSP-dhfr was isolated and ligated. This vector contains pBR322 sequences up to the *Hind*III site and sequences beginning at -780 relative to the initiation site of the *hsp70* promoter. pHSH3-dhfr: pHSP-dhfr was cleaved with *Xho*I, the ends were blunted and *Eco*RI linkers were added. The plasmid was cleaved with *Eco*RI and circularized. The resultant plasmid was cleaved with *Eco*RI and *Hind*III to remove the SV40 origin fragment and the 29-bp *Eco*RI/*Hind*III fragment of pBR322 was ligated in its place. This construction contains sequences to -200 of the *hsp70* promoter fused to pBR322 sequences. pHSH3-cat: Plasmid p232.3⁸ was cleaved with *Pst*I and blunt ends were made. The resultant DNA was cleaved with *Hind*III, the 0.86-kb promoter-containing piece was isolated and cloned between the *Hind*III and *Sma*I sites of p106-cat. This vector contains sequences to -780 of the *hsp70* promoter. p106-cat was a gift of M. Gilman and contains the *Sma*I to *Eco*RI fragment of pSmaI-cat (ref. 12) cloned between the *Sma*I and *Eco*RI sites of the polylinker in pUC13.

cause of this variability is that the *myc* protein appears toxic to CHO cells when expressed constitutively at high levels (R.E.K., unpublished observations). We therefore tested the ability of the *myc* protein to regulate expression of *hsp70*-containing hybrids in a transient transfection protocol. A *Hind*III/*Pst*I fragment of the *hsp70* promoter was linked to the *cat* gene to produce pHSH3-cat (Fig. 1b). This gene was transfected into CHO DUKX B1 cells in the presence or absence of pSV2-cmyc, and cell extracts were made 48 h later. The level of CAT activity in these extracts was determined by incubating the extract with ¹⁴C-chloramphenicol and acetyl-CoA and measuring the amount of acetylated chloramphenicol produced after separation by TLC¹². Co-transfection of pSV2-cmyc with pHSH3-cat into

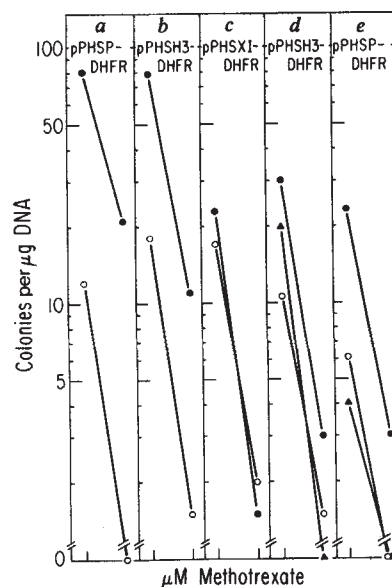


Fig. 2 The effect of co-transfection with pSV2-cmyc on stable colony formation in CHO DUKX B1 cells. The indicated pHS-series plasmid (2 µg) was transfected into CHO DUKX B1 cells with either carrier DNA (○), or with a mixture of carrier DNA and pSV2-cmyc (●) or pfsmyc-20 (▲). In a, 0.4 µg pSV2-cmyc was used, in b-d 2.0 µg pSV2-cmyc or pfsmyc-20 was used and in e 1.0 µg pSV2-cmyc or pfsmyc-20 was used. In all cases, the total amount of DNA transfected was 6 µg, with pBR322 DNA making up the remaining DNA. The data shown are representative examples of experiments repeated at least three times with similar results. In eight separate experiments with pHSP-dhfr, the degree of stimulation caused by co-transfection with pSV2-cmyc was (computed at 0 methotrexate): ×8; ×7; ×6; ×7; ×4; ×4; ×3; ×1. In three separate experiments with pHSH3-dhfr, the degree of stimulation was: ×5; ×4; ×2. The two points on the abscissa in each panel refer respectively to 0 and 0.005 µM methotrexate.

Methods: A detailed description of the stable transfection protocol is presented elsewhere¹⁴. Briefly, CHO DUKX B1 cells were subcultured at a density of 1×10^6 per 10-cm^2 dish 20 h before transfection. The DNA was applied as a calcium phosphate precipitate²⁸, cells were glycerol-shocked (10% glycerol) 4 h after transfection, and then fed with non-selective media (α^- media supplemented with 10% fetal calf serum and with $10 \mu\text{g ml}^{-1}$ each adenosine, deoxyadenosine and thymidine). Two days after transfection, cells were split 1/15 into selective media (α^- with 10% dialysed fetal calf serum) that contained either 0 or 0.005 µM methotrexate, as indicated. Twelve days after transfection, colonies were stained and counted.

CHO DUKX cells stimulated expression of CAT in this assay (data not shown).

We next investigated whether the observed stimulation of *hsp70* hybrid genes would function in cell types other than CHO. Co-transfection of pSV2-cmyc with pHSH3-cat into BALB/c 3T3 cells, which have a low basal level of *myc* expression¹⁷, resulted in a 6–15-fold increase in CAT activity (Fig. 3). Co-transfection with pfsmyc-20 produced no stimulation, implying that the effect seen with pSV2-cmyc was due to the *myc* gene product. To determine whether the stimulation was due to specific sequences found on pHSH3-cat, we tested a *cat* hybrid gene containing the promoter region of one of the other *Drosophila hsp70* genes (pDHSP-cat described in ref. 13). The promoter region of pDHSP-cat is 98% homologous with pHSH3-cat between the transcription start site and position -340, but differs completely beyond this point^{9,18}. CAT expression from pDHSP-cat was not stimulated by co-transfection with pSV2-cmyc (Fig. 3), indicating that the *myc* gene product can regulate expression of pHSH3-cat in BALB/c 3T3 cells in a sequence-dependent manner.

Because the *hsp70* promoter is an extremely weak promoter, we have been unable to determine directly either RNA levels or start sites in the transient co-transfection assays and therefore

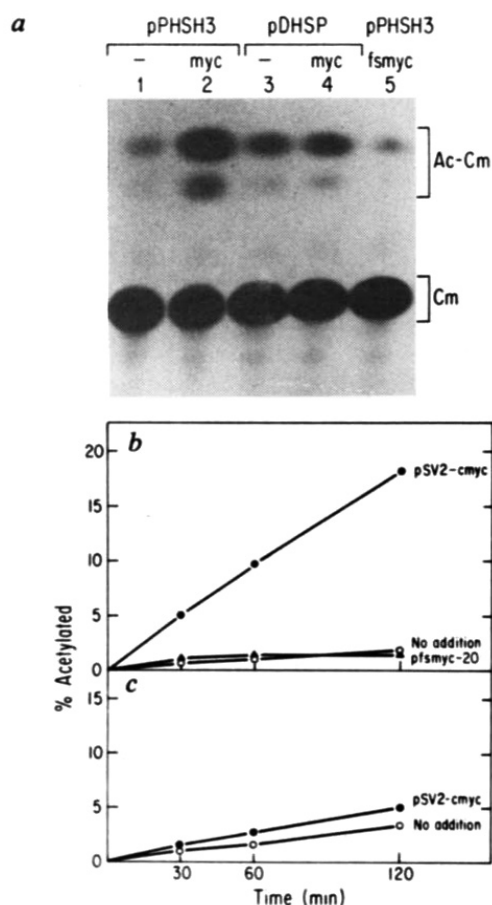


Fig. 3 CAT enzyme activity from lysates of transfected BALB/c 3T3 cells. **a**, Autoradiograph of CAT enzyme assays from cells transfected with plasmids as follows: lane 1, 15 µg of pPHSH3-cat with no addition; lane 2, 15 µg pPHSH3-cat with 1 µg of pSV2-cmyc; lane 3, 15 µg of pDHSP-cat with no addition; lane 4, 15 µg of pDHSP-cat with 1 µg of pSV2-cmyc; lane 5, pPHSH3-cat with 1 µg of pfsmyc-20. The autoradiograph is from the 60-min time point of the enzyme assays. Ac-Cm, acetylated ¹⁴C-chloramphenicol; Cm, ¹⁴C-chloramphenicol. **b**, Results of enzyme assays of cells transfected with 15 µg of pPHSH3-cat alone (no addition), or with 1 µg of either pSV2-cmyc or pfsmyc-20. Results are plotted as per cent of the total ¹⁴C-chloramphenicol acetylated at each time point as indicated. **c**, Results of enzyme assays of cells transfected with 15 µg of pDHSP-cat alone (no addition) or with 1 µg of pSV2-cmyc. These results were generated with lysates from the same transfection experiment. Other identical experiments give 8-, 15-, 6-, 10- and 7-fold stimulation of CAT expression from pPHSH3-cat when co-transfected with pSV2-cmyc.

Methods: Transfections were performed using the CaPO₄ co-precipitation procedure²⁹. Cells were washed with Dulbecco's minimal essential medium (DMEM) and refed DMEM with 10% calf serum 16 h after addition of precipitates. Lysates were prepared ~48 h later by the procedure of Gorman *et al.*¹². 200 µg of protein were assayed from each time point shown. The percentage of acetylated chloramphenicol was determined by cutting out the acetylated and unacetylated ¹⁴C-chloramphenicol regions. Radioactivity was determined by scintillation counting.

cannot conclude that the stimulation observed is at the level of RNA synthesis. The observation that the sequences required for regulation lie more than 200 bases upstream of the normal *hsp70* start site is, however, consistent with this hypothesis. There is evidence for both structural¹⁹ and functional similarities between the *myc* gene product and products of the adenovirus E1a region. Both genes are capable of immortalizing primary cells and of complementing the ability of the c-Ha-ras gene to transform primary cells^{7,19,20}. The E1a region has been shown to stimulate transcription of a wide variety of cellular and viral promoters, including the mammalian *hsp70* gene^{6,14,21-27}. Stimulation of the adenovirus E2 promoter and the human

β-globin promoter by the E1a region does not depend on the presence of a specific regulatory sequence^{14,26}. These results contrast with the experiments described above for *myc* stimulation of the heat shock promoter region, and suggest that the E1a and *myc* gene products may differ in their specificities.

Regardless of mechanism, the observation that the *myc* gene can stimulate gene expression by acting through a specific sequence raises several interesting questions. Does the *myc* gene product bind directly to this sequence, or does it interact with or induce a cellular function that recognizes these sequences? The observation that the *myc* protein is localized to the nucleus is consistent with its ability to bind to DNA sequences^{4,5}. That the mouse *myc* gene can apparently act through *Drosophila* sequences suggests either that the mechanism of stimulation is strongly conserved over evolution or that sequences capable of responding to *myc* protein are fortuitously contained on this particular promoter. This observation then raises the possibility that *myc* can regulate gene expression through similar sequences in mammalian cells, perhaps allowing coordinated regulation of a set of genes.

We thank P. DiNocera, I. Dawid, C. Gorman, M. Gilman, H. Land and L. Parada for donation of plasmid DNAs, R. A. Weinberg for stimulating discussions, M. Gilman for comments on the manuscript and R.E.S.D. Myers for assistance in making plasmid pfsmyc-20. R.E.K. is a fellow of the Jane Coffin Childs Memorial Fund for Medical Research, A.S.B. is a fellow of the NIH. This work was supported by NSF grant PCM-8200309, NIH grants R01-GM32467 and PPG P01-CA26717 to P.A.S. and partially by NIH core grant P01-CA14051.

Received 25 June; accepted 27 September 1984.

- Hayward, W. S., Neel, B. G. & Astrin, S. M. *Nature* **290**, 475-479 (1981).
- Shen-Ong, G. L. C., Keath, E. J., Piccoli, S. P. & Cole, M. D. *Cell* **31**, 443-450 (1982).
- Leder, P. *et al. Science* **222**, 765-771 (1983).
- Donner, P., Greiser-Wilke, I. & Moelling, K. *Nature* **296**, 262-266 (1982).
- Abrams, H. D., Rohrschneider, L. R. & Eisenman, R. N. *Cell* **29**, 427-439 (1982).
- Imperiale, M. J., Kao, H.-T., Feldman, L. T., Nevins, J. R. & Strickland, S. *Molec. cell Biol.* **4**, 867-874 (1984).
- Land, H., Parada, L. F. & Weinberg, R. A. *Nature* **304**, 596-601 (1983).
- Holmgren, R., Livak, K., Morimoto, R., Freund, R. & Meselson, M. *Cell* **18**, 1359-1370 (1979).
- Karch, F., Torok, I. & Tissieres, A. *J. molec. Biol.* **148**, 219-230 (1981).
- Mason, P. J., Torok, I., Kiss, I., Karch, F. & Udvardy, A. *J. molec. Biol.* **156**, 21-35 (1982).
- Kaufman, R. J. & Sharp, P. A. *Molec. cell Biol.* **2**, 1304-1313 (1982).
- Gorman, C. M., Moffat, L. F. & Howard, B. H. *Molec. cell Biol.* **2**, 1044-1051 (1982).
- Di Nocera, P. P. & Dawid, I. B. *Proc. natn. Acad. Sci. U.S.A.* **80**, 7095-7098 (1983).
- Kingston, R. E., Kaufman, R. J. & Sharp, P. A. *Molec. cell Biol.* **4**, 1970-1977 (1984).
- Urlaub, G. & Chasin, L. A. *Proc. natn. Acad. Sci. U.S.A.* **77**, 4216-4220 (1980).
- Urlaub, G., Kas, E., Carothers, A. M. & Chasin, L. A. *Cell* **33**, 405-412 (1983).
- Kelley, K., Cochran, B. H., Stiles, C. D. & Leder, P. *Cell* **35**, 603-610 (1983).
- Hackett, R. W. & Lis, J. T. *Proc. natn. Acad. Sci. U.S.A.* **78**, 6196-6200 (1981).
- Ralston, R. & Bishop, J. M. *Nature* **306**, 803-806 (1983).
- Ruley, H. E. *Nature* **304**, 602-607 (1983).
- Nevins, J. R. *Cell* **29**, 913-919 (1982).
- Berk, A. J., Lee, F., Harrison, T., Williams, J. & Sharp, P. A. *Cell* **17**, 935-944 (1979).
- Bos, J. L. & ten Wolde-Kraamwinkel, H. C. *EMBO J.* **2**, 73-76 (1983).
- Elkaim, R., Coding, C. & Keding, C. *Nucleic Acids Res.* **11**, 7105-7117 (1983).
- Gaynor, R. B., Hillman, D. & Berk, A. J. *Proc. natn. Acad. Sci. U.S.A.* **81**, 1193-1197 (1984).
- Green, M. R., Treisman, R. & Maniatis, T. *Cell* **35**, 137-148 (1983).
- Treisman, R., Green, M. R. & Maniatis, T. *Proc. natn. Acad. Sci. U.S.A.* **80**, 7428-7432 (1983).
- Maxam, A. M. & Gilbert, W. *Proc. natn. Acad. Sci. U.S.A.* **74**, 560-564 (1977).
- Graham, F. L. & van der Eb, A. J. *Virology* **52**, 456-467 (1973).

DNA sequences from the quagga, an extinct member of the horse family

Russell Higuchi*, Barbara Bowman*, Mary Freiburger*, Oliver A. Ryder† & Allan C. Wilson*

* Department of Biochemistry, University of California, Berkeley, California 94720, USA

† Research Department, San Diego Zoo, San Diego, California 92103, USA

To determine whether DNA survives and can be recovered from the remains of extinct creatures, we have examined dried muscle from a museum specimen of the quagga, a zebra-like species (*Equus quagga*) that became extinct in 1883 (ref. 1). We report that DNA was extracted from this tissue in amounts approaching 1% of that expected from fresh muscle, and that the DNA was of

relatively low molecular weight. Among the many clones obtained from the quagga DNA, two containing pieces of mitochondrial DNA (mtDNA) were sequenced. These sequences, comprising 229 nucleotide pairs, differ by 12 base substitutions from the corresponding sequences of mtDNA from a mountain zebra, an extant member of the genus *Equus*. The number, nature and locations of the substitutions imply that there has been little or no postmortem modification of the quagga DNA sequences, and that the two species had a common ancestor 3–4 Myr ago, consistent with fossil evidence concerning the age of the genus *Equus*².

Rau supplied us with a small piece of dried muscle and connective tissue attached to the salt-preserved skin of a quagga (*Equus quagga*) that died 140 yr ago and which had been stored in the Museum of Natural History at Mainz, West Germany³. After releasing the DNA from 0.7 g of this tissue with proteinase K and detergent⁴, we purified it by phenol extraction and ethanol precipitation, then fractionated it by gel-filtration and electrophoresis. Most of the DNA detectable by fluorescence in the presence of ethidium bromide was <500 base pairs (bp) long, if double-stranded. After transfer to nitrocellulose paper⁵, this DNA annealed with total genomic DNA of the mountain zebra (*Equus zebra*) that had been labelled with ³²P by nick-translation⁶. The intensity of the hybridization signal was roughly that expected if the ethidium fluorescence seen in the gel were due to DNA closely related to the probe; by contrast, when human DNA is probed with mountain zebra DNA, 1,000 times more human DNA is needed to give the same signal. These results imply that there is as much as 5 µg of low-molecular weight, zebra-like DNA per gram of the quagga tissue.

About 10 ng of quagga DNA were cloned into the λ gt10 vector, as detailed in Fig. 1 legend; 25,000 resultant plaques containing quagga DNA inserts were screened by hybridization to ³²P-labelled mtDNA from the mountain zebra. Plaques showing a positive reaction were picked off and purified, then subcloned into the *Eco*RI site of phage M13mp11, which provided templates for dideoxy sequencing⁷. (That nuclear DNA was represented among the quagga DNA clones was shown by screening the same plaque lifts with purified *E. zebra* satellite DNA as a probe.) These quagga mtDNA clones were then used

to identify *E. zebra* mtDNA clones containing homologous sequences.

Figure 1 compares the sequences of two quagga clones with the corresponding sequences from *E. zebra* mtDNA. By comparing these sequences with the known sequences of cow⁸ and human⁹ mtDNA, we established that they code for parts of two proteins: 117 bp originating from a gene (unidentified reading frame 1, URF1) coding for a protein of unknown function and 112 bp from the cytochrome oxidase I gene. At the 229 positions compared, the quagga and zebra sequences differ by 12 base substitutions, of which all are transitions and 10 do not cause amino acid replacements. There are no premature termination codons, deletions or additions relative to the zebra sequences. Such a pattern of divergence is close to that expected from previous mtDNA sequence comparisons of other species^{10,11}. A high ratio of transitions to transversions, and of silent changes to amino acid replacements, is typical of mtDNAs from closely related species; the absence of transversions between the equine sequences examined here is not surprising^{10,11}. The degree of sequence difference between the two species (~5%) is in approximate agreement with that estimated by George and Ryder for several pairs of living species of *Equus*, based on restriction maps of the whole mitochondrial genome¹².

A few of the differences between the quagga and zebra DNA sequences could be due to postmortem change (including cloning artefacts). At the 229 positions compared, the cow differs by three amino acid replacements from the zebra and by five from the quagga. Hence, the two amino acid differences (Fig. 1) between the zebra and quagga probably arose in the quagga lineage. A ratio of two changes on the quagga lineage to zero on the zebra lineage to three on the lineage linking the cow with the common ancestor of the quagga and zebra is a little surprising when considered in relation to the durations of these lineages, which are in the ratio of ~1:1:38 (see Fig. 2). Further sequencing of mitochondrial and nuclear DNA from quagga and zebra will permit a better assessment of the likelihood of postmortem change.

The phylogenetic tree in Fig. 2 shows how the quagga mtDNA sequences are related to those of three other mammals. This

Unidentified reading frame 1	
Quagga	C CCA ATC CTG CTC GCC GTA GCA TTC CTC ACA CTA GTT GAA CGA AAA GTC TTA GGC TAC ATA CAA CTT CGT AAA GGA CCC AAC ATC GTA GGC CCC TAT GGC CTA CTA CAA CCC ATT AC
Zebra	T.....G.....T.....C.....G [*]
Cytochrome oxidase I	
Quagga	A GGA GGA TTC GTT CAC TGA TTC CCT CTA TTC TCA GGA TAC ACA CTC AAC CAA ACC TGA GCA AAA ATT CAC TTT ACA ATT ATA TTC GTA GGG GTC AAC ATA ATT TTC TTC CCA
Zebra	G.....T.....G.....C.....A.....T.....C [*]

Fig. 1 Sequences of the coding strands determined for two pieces of quagga mtDNA. The sequences are arranged in triplets corresponding to the amino acids that they encode. At 12 positions, the quagga sequences differ from those of mtDNA from a mountain zebra; only for these positions is the nature of the base specified for the zebra. The two asterisks identify triplets at which the zebra and quagga differ by an amino acid replacement.

Methods: Quagga DNA was cloned by the sequential blunt-ending and repair of the fragments with T4 DNA polymerase and *Escherichia coli* DNA polymerase¹⁹; addition of synthetic linker molecules containing the *Eco*RI restriction endonuclease recognition sequence²⁰; cleavage of the added linker with *Eco*RI to produce sticky ends; and ligation to the *Eco*RI-cleaved DNA of the λ phage cloning vector, λ gt10 (ref. 21) (this vector is suitable for the efficient cloning of small amounts of low-molecular weight DNA). To ~10 ng of quagga DNA were added 200 ng of *Eco*RI-cut λ gt10 and 1.5 µg of λ cI857 Sam7 DNA (New England Biolabs). These DNAs were precipitated with ethanol, dissolved in 4 µl of 10 mM Tris-HCl (pH 7.4), 10 mM MgCl₂, 50 mM NaCl, and kept at 67 °C for 5 min, then placed in a water bath at 50 °C for 15 min; 0.5 µl each of 10 mM ATP, 100 mM dithiothreitol and T4 DNA ligase (New England Biolabs) were added and the reaction incubated overnight at 4 °C. This DNA was then packaged *in vitro* into phage particles²² and allowed to infect *E. coli* BNN102 (ref. 23) having the *hfl150* genotype²⁴. (This host becomes a lysogen so frequently that intact or reassembled λ gt10 genomes form only minute plaques. Only genomes with inserts that interrupt the *cI* gene (λ repressor) will form large, clear plaques. Pre-annealing with λ cI857 Sam7 DNA favours the formation of multimeric genomes (monomeric genomes will not package into phage; ref. 25), which increases the efficiency of the cloning. Because this DNA requires a *supF* host for propagation, it does not form plaques on BNN102.) The resultant 'library' of non-lysogenic phage was screened by annealing ³²P-labelled, purified mountain zebra mtDNA to 'plaque lifts'²⁶ of the library. Plaques showing a positive reaction with this probe were picked off and phage from them purified and the DNAs isolated. The inserts in these phage genomes were removed with *Eco*RI and recloned into the *Eco*RI site of the replicative form of the genome of M13mp11 (Pharmacia P-L Biochemicals); this provided single-stranded M13 DNA templates containing quagga DNA sequences suitable for sequencing via the primed-synthesis, dideoxynucleoside chain-termination method of Sanger *et al.*⁷. Fragments of mountain zebra mtDNA that had been cut with *Taq*I and *Hpa*II were subcloned into the *Acc*I site of M13mp11. Zebra clones containing DNAs related to the two quagga sequences were identified by probing with radioactively labelled quagga clones, then sequenced.

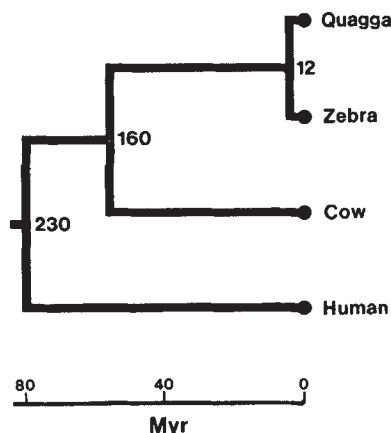


Fig. 2 A phylogenetic tree relating mtDNA sequences from the quagga to those of three other mammals. The branching order of the tree was determined by applying the parsimony method¹¹ to the amino acid sequences read from the nucleotide sequences; this method eliminates the phylogenetic noise created by multiple substitutions at silent sites²⁷. The positions of the nodes in the tree are in proportion to the corrected number of substitutions by which the nucleotide sequences differ. These numbers (12, 160 and 230) were obtained for distantly related sequences by multiplying the observed number of transversions by 10, based on recommendations by Brown *et al.*¹¹; no correction was needed for the quagga-zebra comparison because all the changes were transitions. The cow and human sequences were taken from Anderson *et al.*^{8,9}.

tree accounts for the observed sequence differences with fewer mutational events than trees with other branching orders. The biochemical tree matches that predicted by anatomical evidence, which associates the zebra and quagga most closely with each other, and more closely with the cow than with man. The time scale for the tree is based on the assumption that the time elapsed since the divergence of primate and ungulate lineages is 80 Myr¹¹ and that the rate of divergence among the sequences is constant. Divergence was estimated by correcting the observed sequence differences for multiple hits at the same nucleotide site, as discussed in Fig. 2 legend. The resulting estimates of 3–4 Myr for the quagga-zebra split and 55–60 Myr for the split between the odd-toed and even-toed ungulates are in satisfactory agreement with fossil evidence². This agreement supports the notion that evolution of mtDNA is approximately clock-like.

The present report seems to be the first demonstration that clonable DNA sequence information can be recovered from the remains of an extinct species. The only genetic macromolecules recovered previously from such remains were proteins¹³. By further sequencing of the mitochondrial and nuclear DNAs of the quagga and related extant species, we hope to obtain an evolutionary history of the genus *Equus* in which the quagga is accurately placed; the phylogenetic position of the quagga within this genus is disputed^{14–16}. In a preliminary report of some of the findings presented here¹⁷, we also described the recovery of DNA from a 40,000 yr-old, frozen Siberian mammoth and raised the possibility that DNA occurs in extracts of 25-Myr-old insects preserved in amber¹⁸. If the long-term survival of DNA proves to be a general phenomenon, several fields including palaeontology, evolutionary biology, archaeology and forensic science may benefit.

We thank R. Rau for the quagga tissue and M. George for highly purified mountain zebra DNA; B. Tong for help with the cloning; M. McClelland for help with computer sequence searches; R. Cann and K. Vincent for advice on DNA sequencing; R. T. White for advice on DNA extraction; T. Huynh and R. Davis for λ gt10 and C. Holcomb for advice on its use; and E. Prager for statistical help and editorial advice. NIH and NSF grants provided partial support for this work.

Received 25 June; accepted 6 September 1984.

1. Dolan, J. M. *Zoonoos* **56**, 13–15 (1983).
2. Savage, D. E. & Russell, D. E. *Mammalian Paleofaunas of the World*, 346–397 (Addison-Wesley, Reading, Massachusetts, 1983).
3. Rau, R. *Ann. S. Afr. Mus.* **65**, 41–87 (1974); **77**, 27–45 (1978).
4. Robbins, J. *et al. J. biol. Chem.* **254**, 6187–6195 (1979).
5. Southern, E. M. *J. molec. Biol.* **98**, 503–517 (1975).
6. Rigby, P. W. J., Dieckmann, M., Rhodes, C. & Berg, P. *J. molec. Biol.* **113**, 237–251 (1977).
7. Sanger, F., Nicklen, S. & Coulson, A. R. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5463–5467 (1977).
8. Anderson, S. *et al. J. molec. Biol.* **156**, 683–717 (1982).
9. Anderson, S. *et al. Nature* **290**, 457–465 (1981).
10. Brown, G. G. & Simpson, M. V. *Proc. natn. Acad. Sci. U.S.A.* **79**, 3246–3250 (1982).
11. Brown, W. M., Prager, E. M., Wang, A. & Wilson, A. C. *J. molec. Evol.* **18**, 225–239 (1982).
12. George, M. & Ryder, O. A. *Genetics* **104**, s27 (1983).
13. Prager, E. M., Wilson, A. C., Lowenstein, J. M. & Sarich, V. M. *Science* **209**, 287–289 (1980).
14. Bennett, D. K. *Syst. Zool.* **29**, 272–287 (1980).
15. Eisenmann, V. *Cah. Paléont.* **1**–186 (1980).
16. Groves, C. P. & Willoughby, D. P. *Mammalia* **45**, 321–354 (1981).
17. Higuchi, R. & Wilson, A. C. *Fedn Proc.* **43**, 1557 (1984).
18. Poinar, G. O. & Hess, R. *Science* **215**, 1241–1242 (1982).
19. O'Farrell, P. *Focus* **3**, 1 (1981).
20. Scheller, R. H., Dickerson, R. E., Boyer, H. W., Riggs, A. D. & Itakura, K. *Science* **196**, 177–180 (1977).
21. Huynh, T. V., Young, R. A. & Davis, R. W. in *DNA Cloning Techniques: A Practical Approach* (ed. Glover, D. M.) (IRL, Oxford, 1984).
22. Hohn, B. *Meth. Enzym.* **68**, 299–309 (1979).
23. Young, R. A. & Davis, R. W. *Proc. natn. Acad. Sci. U.S.A.* **80**, 1194–1198 (1983).
24. Hoyt, M. A., Knight, D. M., Das, A., Miller, H. I. & Echols, H. *Cell* **31**, 565–573 (1982).
25. Feiss, M. & Becker, A. in *Lambda II* (eds Hendrix, R. W., Roberts, J. W., Stahl, F. W. & Weisberg, R. A.) 305–330 (Cold Spring Harbor Laboratory, New York, 1983).
26. Woo, S. L. C. *Meth. Enzym.* **68**, 389–395 (1979).
27. Doolittle, R. F. *Science* **214**, 149–159 (1981).

Nitrogen fixation by a methanogenic archaeobacterium

Patti A. Murray & Stephen H. Zinder

Department of Microbiology, New York State College of Agriculture and Life Sciences, Ithaca, New York 14853, USA

The ability to fix nitrogen (N_2) is found among a wide variety of the prokaryotic eubacteria, but not in eukaryotes¹. In addition to the prokaryotic eubacteria and eukaryotes, a third 'kingdom'—the archaeobacteria—has been defined based on the comparison of 16S ribosomal oligonucleotide sequence catalogues^{2,3}. Included in the archaeobacterial kingdom are certain obligate halophiles and thermoacidophiles, and the methanogens, strictly anaerobic, methane-producing bacteria⁴. Here we report diazotrophy by an archaeobacterium, the methanogen *Methanosarcina barkeri* strain 227. Because it has been proposed that the archaeobacteria, eubacteria and eukaryotes diverged at an early stage in evolution^{2,3}, the discovery of diazotrophy (N_2 fixation) in a member of the archaeobacterial group raises interesting evolutionary questions.¹

In the thermophilic *Methanosarcina* strain TM-1 (ref. 5), growth and methanogenesis can be limited by 1 mM NH_4^+ as the sole source of fixed nitrogen in the growth medium⁶, as has been found in other methanogens^{7,8}. Under such nitrogen-limiting conditions, *Methanosarcina* strain TM-1 stores a glycogen-like polysaccharide in amounts up to 20 mg per g protein⁶. We initiated a similar comparative study using a mesophile, *Methanosarcina barkeri* strain 227 (ref. 9). However, this culture showed no signs of nitrogen limitation of growth, and could be transferred repeatedly in methanol growth medium from which NH_4^+ was omitted. The headspaces in these culture vials consisted of 70% N_2 /30% CO_2 . To test whether N_2 was the nitrogen source of these cultures, an N_2 -free argon headspace was used. Figure 1a shows that, in the absence of both NH_4^+ and N_2 , methanogenesis by *M. barkeri* strain 227 was significantly decreased. When 14% N_2 was added to the argon headspace, there was nearly the same final yield of methane as in the vials provided with 20 mM NH_4^+ . Cell growth, measured as dry weight, was only 15 μg ml⁻¹ on day 16 in vials having only argon as the headspace, but was 95 μg ml⁻¹ in vials with N_2 added.

This apparent ability of the *Methanosarcina* culture to fix N_2 could possibly be caused by an N_2 -fixing contaminant. However,

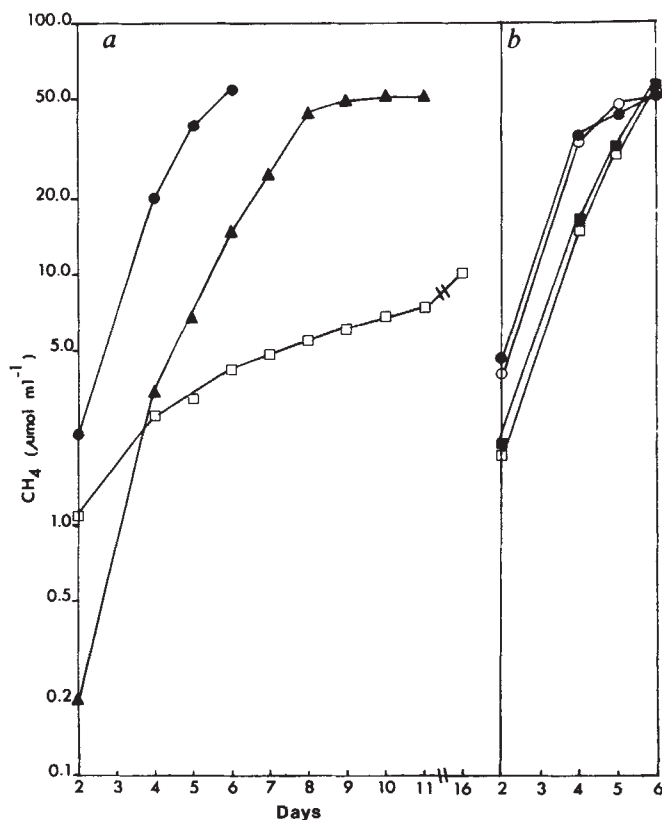


Fig. 1 *a*, Production of CH_4 by *Methanosarcina barkeri* strain 227 under an argon headspace with 20 mM NH_4^+ (●), 14% N_2 (▲), or no nitrogen source added (□). *b*, Production of CH_4 by *M. barkeri* strain 227 under a 70% N_2 /30% CO_2 headspace with 20 mM added NH_4^+ (○), 20 mM NH_4^+ and 100 $\mu\text{g ml}^{-1}$ vancomycin (●), no NH_4^+ (□), no NH_4^+ and 100 $\mu\text{g ml}^{-1}$ vancomycin (■).

Methods: *a*, Cells were grown in a medium containing (g per l of distilled water): K_2HPO_4 , 0.34; $\text{MgCl}_2 \cdot 6\text{H}_2\text{O}$, 0.1; resazurin, 0.001; trace metals solution¹⁶ (with the addition of 0.02 g $\text{NiCl}_2 \cdot 6\text{H}_2\text{O}$ per l), 10 ml l^{-1} . The pH of the medium was adjusted to 6.7 with 1.2 M HCl, and the medium was then boiled under argon. The medium was dispensed in 50-ml quantities into 118-ml vials sealed with butyl rubber stoppers in an anaerobic glove box and autoclaved. After autoclaving, each vial was flushed repeatedly with sterile argon scrubbed with hot copper coils. The following were added as sterile, anaerobic solutions to each vial: 72 mM methanol; 0.01% $\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$; 0.04% $\text{Na}_2\text{S} \cdot 9\text{H}_2\text{O}$; 24 mM NaHCO_3 followed by 18 mM HCl to create 30% CO_2 in the headspace. 20 mM NH_4Cl or 10 ml N_2 were added when appropriate. Culture vials received a 1% inoculum of methanol-grown cells previously transferred several times in medium with no NH_4Cl , under a 70% N_2 /30% CO_2 headspace. (The only other added potential nitrogen source available in this medium, other than N_2 , was 0.2 mM nitrilotriacetic acid from the trace metal solution.) Methane was quantified by gas chromatography as outlined by Zinder *et al.*¹⁷. High purity argon (minimum 99.995% pure) was from Matheson Gas Products. *b*, Each 50-ml growth vial was prepared as in *a*, except that the NaHCO_3 added to each vial was reduced to 6 mM, no HCl was added, and vials were flushed aseptically with 70% N_2 /30% CO_2 . In addition, 100 $\mu\text{g ml}^{-1}$ of anaerobic, filter-sterilized vancomycin (Sigma) was added to the appropriate vials.

repeated microscopic examinations revealed no contaminants and no growth of contaminants was detected in routine growth medium (Fig. 1) supplemented with 0.1% Difco yeast extract and 0.05% sugars (glucose, sucrose, xylose, cellobiose) added or in modified brewer's thioglycollate medium (Fisher Scientific). To test for the presence of contaminating sulphate-reducing bacteria, known to be diazotrophic¹, we inoculated a culture of *M. barkeri* strain 227 grown under N_2 -fixing conditions into NH_4^+ -containing routine growth medium supple-

Table 1 Final A_{600} , cell protein and cell dry weight of *Methanosarcina barkeri* strain 227 grown under an N_2 / CO_2 headspace in the presence or absence of NH_4^+ and vancomycin

Growth conditions	Final	Final protein ($\mu\text{g ml}^{-1}$)	Final dry weight ($\mu\text{g ml}^{-1}$)
+ NH_4^+			
–Vancomycin	0.65	42	150
+ NH_4^+			
+Vancomycin	0.6	45	100
– NH_4^+			
–Vancomycin	0.66	46	100
– NH_4^+			
+Vancomycin	0.7	48	120

Cells were grown as described in Fig. 1*b* legend. To measure final A_{600} , 10 ml samples of cells were removed at the end of exponential growth (day 6 in Fig. 1*b*). The resazurin in these samples was reduced using sodium hydrosulphite. A_{600} was measured in 18-mm test tubes using a Sequoia-Turner model 340 spectrophotometer. Two 5-ml samples of cells were removed at the same time for protein analysis and dry weight measurements. For the protein determinations, the cells were centrifuged, washed in 0.1 M potassium phosphate buffer, pH 6.7, resuspended in 1.0 ml of buffer and broken. The cells were broken by shaking with ~1 ml of 0.45-mm glass beads in a screw-capped tube at maximum speed using a Vortex mixer. The protein concentration was measured using the Coomassie brilliant blue method¹⁴ (reagents from Bio-Rad). Cellular dry weight was determined by filtering 5 ml of cells through a pre-dried, pre-weighed 0.2 μm Nuclepore membrane filter and drying in a desiccator at 50 °C for 24 h.

mented with 10 mM lactate and sulphate, as well as a vitamin solution⁴ and 0.2 g l^{-1} yeast extract. No growth could be detected in this medium after incubation for 7 days at 37 °C, while another vial of this medium inoculated with a small amount of anaerobic pond sediment was found to have become noticeably blackened and to contain many motile spiral-shaped bacteria similar in morphology to *Desulfovibrio*. Also, no growth other than *M. barkeri* strain 227 was obtained when inoculated into sulphate-containing medium in which either methanol or acetate replaced lactate as the electron donor.

Another demonstration of culture purity was that growth of strain 227 under N_2 -fixing or non-fixing conditions was not affected by the antibiotic vancomycin (Fig. 1*b*). Vancomycin is a strong inhibitor of peptidoglycan synthesis in eubacteria and should inhibit growth of a putative N_2 -fixing contaminant, but should not affect the archaeobacterium, *M. barkeri* strain 227. When strain 227 was grown in the presence of vancomycin, growth under N_2 -fixing conditions was only slightly slower than growth with NH_4^+ . In this experiment, culture absorbance at 600 nm (A_{600}), cellular protein levels and dry weights were measured on 5- or 10-ml culture samples taken on day 6 (Table 1). The presence of the antibiotic did not affect the yields. In addition, cultures under N_2 and non- N_2 -fixing conditions showed similar A_{600} , protein and dry weight values. Because these cultures were collected after growth had ceased, the dry weight yield values obtained may not represent growth yields, which should be obtained for exponentially growing cultures. Detailed studies are in progress to determine the effect of N_2 fixation, believed to be a metabolically costly process¹, on growth yields of *M. barkeri* strain 227.

M. barkeri strain 227 also grew under N_2 -fixing conditions with acetate or H_2 / CO_2 as the methanogenic substrate, but growth was considerably slower than with NH_4^+ as the nitrogen source (data not shown). In the case of H_2 / CO_2 , this may be due to the ability of H_2 to inhibit nitrogenase¹.

The N_2 -fixing capability of *M. barkeri* strain 227 was demonstrated unequivocally by $^{15}\text{N}_2$ incorporation. Acetylene reduction to ethylene is often used to show N_2 fixation¹, but the technique is not suitable in this case because acetylene is a potent inhibitor of methanogenesis¹⁰, and because it is conceivable that methanogens could reduce acetylene without

Table 2 $^{15}\text{N}_2$ incorporation by *Methanosarcina barkeri* strain 227

Growth conditions	^{15}N (% total cellular N)
$^{15}\text{N}_2$, $-\text{NH}_4^+$	17.8 ± 2.6
$^{14}\text{N}_2$, $-\text{NH}_4^+$	0.33 ± 0.0
$^{15}\text{N}_2$, $+\text{NH}_4^+$	0.36 ± 0.02

The medium and the 50 ml culture vials were prepared as outlined in Fig. 1b, but 240 mM methanol was provided (methanol was added 48 mM at a time) and the vancomycin was absent. Before incubation, 20 ml of sterile, anaerobic $^{15}\text{N}_2$ (99.9% $^{15}\text{N}_2$, Monsanto) was added to the 70% N_2 /30% CO_2 headspace of replicate vials. The cells were collected after depletion of the methanol. Cell pellets were washed in 0.1 M potassium phosphate buffer, pH 6.7. The cell pellets were dried, the dried preparations were digested in 18M sulphuric acid, and cellular NH_4^+ was collected by steam distillation into boric acid, both following the Kjeldahl method¹¹. The distillates were titrated, acidified to pH 1.5 with H_2SO_4 , and dried at 75 °C for several days, resuspended in distilled water and redried. The dry material was analysed using a Micro Mass 622 mass spectrometer as outlined by Porter and O'Deen¹⁵.

nitrogenase. When cells were grown in the presence of approximately 30% $^{15}\text{N}_2$, cellular nitrogen was significantly enriched with ^{15}N (Table 2), whereas cells grown without added $^{15}\text{N}_2$ showed natural ^{15}N abundance. Natural ^{15}N abundance was also found in cells grown with $^{15}\text{N}_2$ in the presence of 20 mM NH_4^+ , indicating that these cells do not fix N_2 in the presence of NH_4^+ . In all three cases the cellular nitrogen content (by Kjeldahl analysis¹¹) was ~9% of the total dry weight.

The discovery of diazotrophy in *M. barkeri* strain 227 extends the list of N_2 -fixing organisms to include a member of the archaeobacteria, with interesting implications for the controversy as to the origin and evolution of biological diazotrophy¹. Another methanogen, *Methanococcus thermolithotrophicus*, has been found to fix N_2 when growing on H_2/CO_2 as the methanogenic substrate (see accompanying paper¹²). The archaeobacteria have evolved separately from the eubacteria, so it will be interesting to establish whether nitrogen fixation in the methanogens is significantly divergent from eubacterial systems or highly conserved, as has been found in the prokaryotic eubacteria¹³. In practical terms, the existence of N_2 -fixing methanogens introduces the possibility of N_2 -fixing microbial systems able to convert wastes which are low in combined nitrogen, including lignocellulose biomass substrates, industrial wastes and food process wastes such as whey, to CH_4 .

We thank Theodore Klein for assistance with the Kjeldahl digestions and Richard Seegers for performing mass spectrometer analysis. This research was funded in part by USDA/Hatch funds and by the Cornell University Graduate Fellowship Office.

Received 18 June; accepted 6 September 1984.

1. Postgate, J. R. *The Fundamentals of Nitrogen Fixation* (Cambridge University Press, 1982).
2. Woese, C. R., Magrum, L. J. & Fox, G. E. *J. molec. Evol.* **11**, 245-252 (1978).
3. Fox, G. E., Magrum, L. J., Balch, W. E., Wolfe, R. S. & Woese, C. R. *Proc. natn. Acad. Sci. U.S.A.* **74**, 4537-4541 (1977).
4. Balch, W. E., Fox, G. E., Magrum, L. J., Woese, C. R. & Wolfe, R. S. *Microbiol. Rev.* **43**, 260-296 (1979).
5. Zinder, S. H. & Mah, R. A. *Appl. env. Microbiol.* **38**, 996-1008 (1979).
6. Murray, P. A. & Zinder, S. H. *83rd An. Meet. Am. Soc. Microbiol.*, Abstr. (American Society for Microbiology, Washington, DC, 1983).
7. Kenaley, W. R., Thompson, T. E., Schubert, K. R. & Zeikus, J. G. *J. Bact.* **150**, 1357-1365 (1982).
8. Bryant, M. P., Tzeng, S. F., Robinson, I. M. & Joyner, A. E. *Adv. Chem.* **105**, 23-40 (1971).
9. Mah, R. A., Smith, M. R. & Baresi, L. *Appl. env. Microbiol.* **35**, 1174-1184 (1978).
10. Spratt, G. D., Jarrell, K. F., Shaw, K. M. & Knowles, R. J. *gen. Microbiol.* **128**, 2453-2462 (1982).
11. Bremner, J. M. in *Methods of Soil Analysis Pt 2* (eds Black, C. A. et al.) 1256-1286 (American Society of Agronomy, Madison, 1965).
12. Belay, N., Sparling, R. & Daniels, C. *Nature* **312**, 286-288 (1984).
13. Ruvkun, G. B. & Ausubel, F. M. *Proc. natn. Acad. Sci. U.S.A.* **77**, 191-195 (1980).
14. Bradford, M. M. *Analyt. Biochem.* **72**, 248-254 (1976).
15. Porter, L. K. & O'Deen, W. A. *Analyt. Chem.* **49**, 514-516 (1977).
16. Zeikus, J. G. *Bact. Rev.* **41**, 514-541 (1977).
17. Zinder, S. H., Cardwell, S. C., Anguish, T., Lee, M. & Koch, M. *Appl. env. Microbiol.* **47**, 796-807 (1984).

Dinitrogen fixation by a thermophilic methanogenic bacterium

Negash Belay, Richard Sparling & Lacy Daniels

Department of Microbiology, University of Iowa, Iowa City, Iowa 52242, USA

Methanogenic bacteria are known to use NH_4^+ as a nitrogen source for growth¹. Previous work with an impure methanogenic culture suggested that a methanogen might fix atmospheric dinitrogen as a nitrogen source², but no further work on this phenomenon has been documented. We have now examined the use of N_2 by *Methanococcus thermolithotrophicus* and find that the organism can grow well, with multiple transfers, in medium having N_2 as the source of nitrogen. Control cultures without N_2 and containing less than 0.1 mM NH_4^+ do not grow. Growth yields with N_2 are on the average one-third those with NH_4^+ , suggesting that, as in other nitrogen-fixing organisms, this bacterium requires a large amount of ATP for the reduction to occur. After growing in NH_4^+ -containing medium, a long lag is observed before growth begins with N_2 as the nitrogen source; the NH_4^+ levels must be very low for growth to begin. Cells grown in N_2 -fixing conditions reduce acetylene to ethylene. The discovery of a nitrogen-fixing archaeobacterium has important implications for studies on the evolution of nitrogenase, and the fact that *M. thermolithotrophicus* nitrogenase is active at 64 °C suggests that a novel enzyme is involved.

M. thermolithotrophicus (kindly provided by K. O. Stetter³) was grown on H_2/CO_2 in a medium previously described⁴, but containing in addition 0.2 μM tungstic acid and 0.2 μM sodium selenate (see legend to Fig. 1). When NH_4^+ was the nitrogen source the concentration was 16.8 mM. The anaerobic techniques described by Balch and Wolfe⁵ were used to grow liquid cultures with H_2/CO_2 (80:20 v/v) at 64 °C; the technique includes pressurizing bottles and tubes of anaerobic medium to a total of 200-350 kPa of pressure of H_2/CO_2 (80:20 v/v), and repressurizing at intervals to replace the gas consumed. Growth was measured by absorbance at 600 nm in a Perkin Elmer 552A spectrophotometer, and protein was determined by the Bradford assay⁶. Methane production and acetylene reduction to ethylene was measured by gas chromatography (Varian aerograph 2700).

Isolated colonies of *M. thermolithotrophicus* were obtained on Gelrite plates prepared with the regular medium but containing 0.5% Gelrite⁷. The medium was prepared anaerobically in bottles, autoclaved and poured into Petri plates in an anaerobic glove bag. After inoculation of cells onto the plates by spreading or streaking, the plates were placed in a modified pressure cooker⁵. Hydrogen sulphide gas was added from a tank of H_2S via syringe through a rubber stopper in the pressure cooker, making a final H_2S concentration of 1.0%, a concentration close to those suggested for optimal growth on plates of other *Methanococcus* species by Jones et al.⁸. Plates were incubated with a 105 kPa overpressure of H_2/CO_2 at 50 °C. Few colonies were obtained on agar plates, but on Gelrite plates, *M. thermolithotrophicus* formed colonies with greater efficiency than on agar. Isolated colonies were restreaked on plates of identical medium and individual colonies picked for liquid growth experiments. For preparation of inocula for experiments demonstrating N_2 fixation, a 50 ml medium lacking NH_4^+ was prepared in a 250 ml pressure bottle (Wheaton Scientific). This bottle was inoculated with 4 ml of culture grown in the presence of 16.8 mM NH_4^+ (resulting in a carry-over NH_4^+ of 1.2 mM). Further transfers for experiments on the use of NH_4^+ or N_2 were made from this bottle. A culture that used N_2 as the nitrogen source was originally obtained by transferring the low ammonium inoculum (1.2 mM) into 250-ml bottles containing 50 ml of ammonium-free medium. Before inoculation, the bottles were gassed thoroughly (with entry and exit needles) with a 95% N_2 /5% CO_2 gas mixture at 35 kPa above atmos-

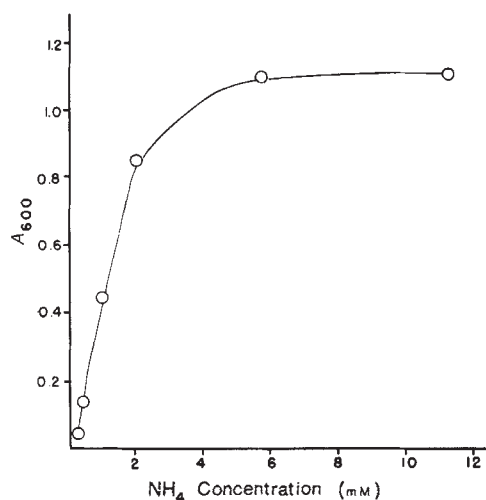


Fig. 1 Dependence of *M. thermolithotrophicus* on ammonium for growth. The absorbance values represent final values. The medium consisted of the following components (mM) in distilled and deionized water: KCl (4.1), $\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$ (0.61), K_2HPO_4 (1.45), KH_2PO_4 (1.85), NaCl (500), $\text{MgCl}_2 \cdot 6\text{H}_2\text{O}$ (38), Na_2CO_3 (1.5), resazurin (0.003), tungstic acid (0.0002), sodium selenate (0.0002). Mg salts were added separately after autoclaving. A 100 $\times$ concentrated solution of trace elements was added to this medium to give the following final concentrations (μM): nitrilotriacetic acid (71), $\text{MnCl}_2 \cdot 4\text{H}_2\text{O}$ (4.5), $\text{FeCl}_2 \cdot 4\text{H}_2\text{O}$ (6.8), CaCl_2 (4.1), $\text{CoCl}_2 \cdot 6\text{H}_2\text{O}$ (7.6), ZnCl_2 (6.6), CuSO_4 (2.8), $\text{Na}_2\text{MoO}_4 \cdot 2\text{H}_2\text{O}$ (1.9), $\text{NiCl}_2 \cdot 6\text{H}_2\text{O}$ (3.8). The pH of the medium was adjusted to 7.6 with Na_2CO_3 before making the medium anaerobic with H_2/CO_2 (80:20, v/v); the final pH was about 6. NH_4Cl was added to give the desired NH_4^+ levels. All data points represent the average of triplicate tubes.

pheric (1 atm = 100 kPa); sterile sodium sulphide (0.5 ml of a 100 mM solution) was re-added to replace the H_2S lost. The bottles were then pressurized to a final overpressure of 140 kPa with H_2/CO_2 (80:20, v/v) (thus there was about 140 kPa of N_2 and 105 kPa of H_2/CO_2 total, in the vessel), and then were inoculated. After approximately half-maximal growth, the bottles were repressurized with H_2/CO_2 . All gases for the gas mixtures used in these experiments (Air Products and Chemicals Inc.) were of research grade (99.9995% pure; maximum guaranteed analysis of $\text{N}_2\text{O} < 0.1$ molar p.p.m.) except for acetylene which was made from calcium carbide. Using cells originally grown with NH_4^+ , there was a long lag (2–3 weeks) before growth on N_2 as the nitrogen source occurred. There was no growth in the control bottles lacking N_2 in the gas phase or NH_4^+ above 0.1 mM. From the first culture growing with N_2 present, several more transfers were made into medium having N_2 as the nitrogen source before conducting methanogenesis and growth yield experiments. The cultures grew faster, and with shorter lag, after more transfers. When a N_2 -fixing culture was plated onto Gelrite Petri plates lacking NH_4^+ and incubated in a $\text{N}_2/\text{H}_2/\text{CO}_2$ atmosphere (pressures as described above), methanogenesis occurred and colonies of the bacteria appeared after 1 week at 50 °C.

The effect of different ammonium levels on growth when NH_4^+ as the nitrogen source was studied. As shown in Fig. 1, no growth was observed below 0.3 mM. Optimal growth occurred above 5 mM, and no inhibition was observed up to 23 mM.

Figure 2a,b shows growth of N_2 adapted cells in medium containing either NH_4^+ or N_2 as the sole nitrogen source. The control cultures without N_2 or NH_4^+ did not grow. When an inoculum from ammonium-grown cells is inoculated into ammonium medium, cells grow without a significant lag; inoculation of N_2 -grown cells into ammonium medium resulted in a 2-h lag, followed by a growth rate equal to the ammonium-grown cells. Cells from the N_2 -adapted culture transferred to ammonium-free N_2 -containing medium grew (after a 2-h lag) originally at the same rate as the NH_4^+ culture, but later the

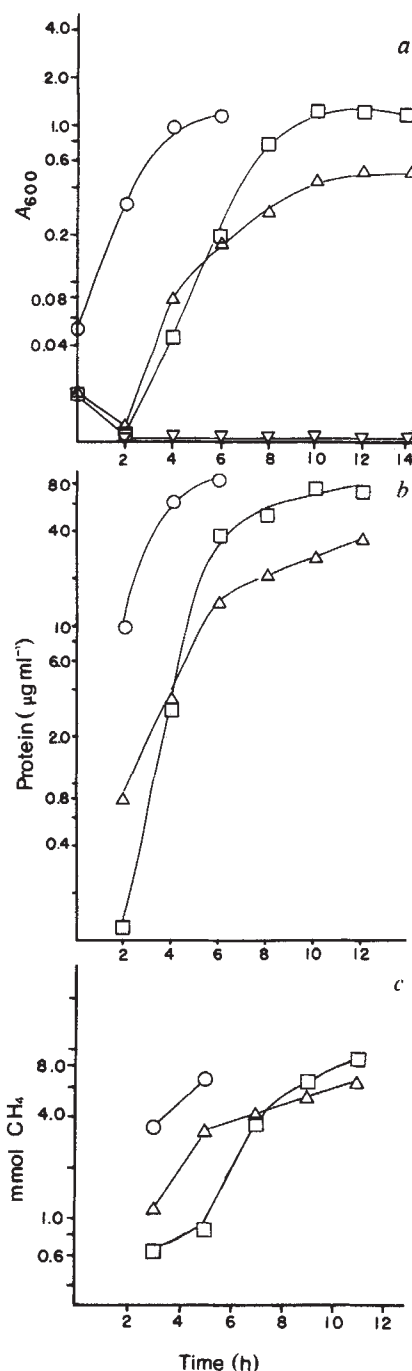


Fig. 2 Growth of *M. thermolithotrophicus* in medium containing either ammonium or dinitrogen as nitrogen source. Data collected from an ammonium-grown culture inoculated from an ammonium culture (○); data from an ammonium-grown culture inoculated from a N_2 -fixing culture (□); data from a culture grown on N_2 as its sole nitrogen source inoculated from a N_2 -fixing culture (Δ); and data from a nitrogen source free culture under $\text{Ar}/\text{H}_2/\text{CO}_2$, inoculated from a N_2 -fixing culture (▽). a, Absorbance of the culture at 600 nm. b, Protein levels of the same cultures. c, Methanogenesis, monitored as total methane per bottle; each bottle (Wheaton Scientific no. 223952) contained 30 ml of medium prepared as described in the text, in a total volume of 540 ml per bottle. All data points represent the average of duplicates.

growth rate slowed. This experiment conclusively shows that growth occurs when NH_4^+ is replaced by N_2 as the nitrogen source, that is that *M. thermolithotrophicus* fixes dinitrogen.

The amount of methane produced by these cultures roughly parallels growth (Fig. 2c). The same conclusion, that dinitrogen fixation occurs, is reached from the methane production data.

Table 1 Growth characteristics of *M. thermolithotrophicus* with ammonium or dinitrogen as nitrogen source

Nitrogen source	Inoculum history	Protein* ($\mu\text{g ml}^{-1}$)	Yield† (g dry cell weight per mol CH_4)	Specific activity‡ ($\mu\text{mol CH}_4$ per min per mg)
None	N_2 or NH_4	—	—	—
NH_4^+	NH_4 grown	89.3	1.54	16.8
NH_4^+	N_2 grown	78.0	0.94	20.2
N_2	N_2 grown	36.4	0.36	11.5

Growth conditions are given in Fig. 2 legend. The data are averages of duplicate experiments. —, Not detectable.

* Maximal protein concentration obtained during growth.

† Yield as measured at the end of log phase. Yields were calculated as the difference of protein concentrations divided by that of methane concentration over the period of logarithmic growth. To obtain gram dry weights it was assumed that proteins comprise 50% of the cells' dry weight.

‡ The rate of methanogenesis during the logarithmic phase of growth, over a 2-h period, the protein concentrations used being the average concentrations over this time period.

The amount of methane produced per cell material produced (the yields in terms of g cell dry weight per mol CH_4) are distinctly different when grown with N_2 as the nitrogen source compared with NH_4^+ . Table 1 describes the cell yield data obtained from the experiment shown in Fig. 2. Both N_2 - and NH_4^+ -growth cells inoculated into NH_4^+ -containing medium show roughly the same yield. The N_2 -fixing experimental bottles gave yields of 0.36 g cell dry weight per mol CH_4 (assuming dry weight is equal to $2 \times$ g protein), about, on average, one-third that of the ammonium-grown cultures. This lower yield is expected with N_2 fixation, as the nitrogenase reaction consumes 18–24 mols ATP per mol N_2 fixed⁹, and thus the cells must make more CH_4 to provide this energy. The values reported here can be compared with earlier methanogen yield data¹⁰ which vary from 0.6 to 3 g cell dry weight per mole CH_4 for H_2/CO_2 growth *Methanobacterium* sp.

The methanogens are known to be inhibited by low levels of acetylene¹¹. Lowered ATP levels rapidly result from exposure of cells to about 1% or more acetylene; thus, the traditional acetylene reduction assay for nitrogenase (using 10% acetylene)¹² would not be expected to work, and indeed our results show no evidence of ethylene production in these conditions; cell cultures to which 10% acetylene was added lysed within 30 min. However, if a lower level of acetylene is used, so as to allow some ATP synthesis, the production of ethylene from acetylene is observed both with N_2 -grown bottle cultures or concentrated, resuspended fermentor grown cells. For example, a 12-l culture was grown in a New Brunswick Scientific Microgen fermentor using formate as the energy source (with medium as described in Fig. 1 legend, except that 200 mM formate and 1 μM sodium selenate and 1 μM tungstic acid were present; no NH_4^+ was present and the gas phase was N_2/CO_2 , 50:50 v/v; pH was continually kept at 6.5–6.6 with formic acid). Collected cells (20% of the total fermentor collection) were resuspended in bottles (total volume 75 ml) with 30 ml of fresh formate medium. Acetylene was added to a final level of 0.3%. Cells grown in N_2 -fixing conditions (no added NH_4^+) produced 15 nmol ethylene per h per mg protein; $\sim 1,000$ mol methane were produced per mol acetylene reduced (see Table 2). However, the system was suboptimal and formate-limited, and thus represents a less than maximal rate. In NH_4^+ growth conditions, no acetylene was reduced by bottle- or fermentor-grown cultures. These data provide further strong evidence for nitrogen fixation.

Concurrently with our observations, Murray and Zinder have found that another methanogen, *Methanosarcina barkeri*, fixes dinitrogen (see accompanying paper¹³). These reports are the first demonstration of nitrogen fixation by a pure culture of

Table 2 Efficiency of growth on various nitrogen sources

Nitrogen source during growth	CH_4 production ($\mu\text{mol per h per mg protein}$)	Acetylene present ($\mu\text{mol per ml gas phase}$)	Ethylene production (nmol per h per mg protein)
N_2	12.5	0.023	15.1
N_2	13.7	0	0
NH_4^+	10.2	0.016	0

Values are the average of duplicates.

methanogenic bacteria, and the first reports for any of the archaeobacteria. The novelty of our finding prompts several questions regarding the evolution of nitrogenase, as archaeobacteria are thought to have separated evolutionarily from eubacterial ancestors very early in evolution. *M. thermolithotrophicus* is also the first thermophile demonstrated to fix dinitrogen at 60 °C or above. The most thermophilic of the N_2 -fixing cyanobacteria, *Mastigocladus*, can grow at temperatures as high as 60–64 °C, but nitrogen fixation probably does not occur at 60 °C or above¹⁴.

Ecologically, the fixation of nitrogen by methanogens has many implications. In anaerobic digestion systems, such as cellulosic conversion to methane, nitrogen could be rate-limiting for the growth of organisms. It is known that some clostridia and other heterotrophs are able to fix N_2 ; these organisms may provide some input of nitrogen into nitrogen-poor anaerobic ecosystems. Now it appears that at least one methanogen has a similar ability. *M. thermolithotrophicus* was originally isolated from a near-shore submarine hydrothermal vent³, which may be low in ammonium nitrogen, and N_2 fixation may be of importance to its ability to live in that environment. Seawater contains virtually no NO_3^- or NH_4^+ , but has significant amounts of N_2 (about 8–16 mg l^{-1} at 10 °C)^{15,16}. In freshwater aquatic ecosystems, nitrate and ammonium levels are typically very low (0.005–1.0 mg l^{-1}) compared with dinitrogen ($\sim 16 \text{ mg l}^{-1}$)¹⁵. Other nitrogen-poor environments where dinitrogen-fixing methanogens might be found are terrestrial hot springs and lakes containing geologically generated H_2 but with little organic nitrogen input.

This work was supported by grants from the NIH (GM30868), the NSF (PCM 82-07809) and the Petroleum Research Fund of the American Chemical Society (14062-AC5). R.S. was supported by a postgraduate scholarship from the Natural Sciences and Engineering Council of Canada.

Note added in proof: Methanogen DNA has been shown to hybridize with *Klebsiella pneumoniae nif KDH* gene probes¹⁷.

Received 13 June; accepted 6 September 1984.

- Bryant, M. P., Tzeng, S. F., Robinson, I. M. & Joyner, A. E. *Adv. Chem.* **105**, 23–40 (1971).
- Pine, M. J. & Barker, H. A. *J. Bact.* **68**, 589–591 (1974).
- Huber, J. M., Thomm, M., Konig, H., Thies, G. & Stetter, K. O. *Archs. Microbiol.* **132**, 47–50 (1982).
- Daniels, L., Belay, N. & Mukhopadhyay, B. *Biotech. Bioengng* (in the press).
- Balch, W. E. & Wolfe, R. S. *Appl. Envir. Microbiol.* **32**, 781–791 (1976).
- Bradford, M. M. *Analyt. Biochem.* **72**, 248–254 (1976).
- Lin, C. C. & Casida, L. E. *Appl. Envir. Microbiol.* **47**, 427–429 (1984).
- Jones, W. J., Whitman, W. B., Fields, R. D. & Wolfe, R. S. *Appl. Envir. Microbiol.* **46**, 220–226 (1983).
- Davis, L. C. & Orme-Johnson, W. H. *Biochim. Biophys. Acta* **452**, 42–58 (1976).
- Daniels, L., Sparling, R. & Sprott, G. D. *Biochim. Biophys. Acta* **768**, 113–163 (1984).
- Sprott, G. D., Jarrell, K. F., Shaw, K. M. & Knowles, R. J. *Gen. Microbiol.* **128**, 2453–2462 (1982).
- Guth, J. H. & Burris, R. H. *Biochemistry* **22**, 5111–5122 (1983).
- Murray, P. A. & Zinder, S. H. *Nature* **312**, 284–286 (1984).
- Brock, T. D. *Thermophilic Microorganisms and Life at High Temperature* (ed. Starr, M. P.) 222 (Springer, New York, 1978).
- Brock, T. D. *Principals of Microbiol. Ecology*, 6 (Prentice-Hall, Englewood Cliffs, 1966).
- Raymont, J. E. G. *Plankton and Productivity in the Oceans*, 33–34 (Pergamon, New York, 1963).
- Pariot, D., Sibold, L., Bhatnagar, L., Henriquet, M. & Aubert, J. P. *Proc. First Symp. Biotechnol. Advances in Processing Municipal Wastes for Fuels and Chemicals* (in the press).

The pollen record and Easter Island statues

THE recent pollen study of Easter Island by Flenley and King¹ is very welcome. However, we must express reservations about their tentative conclusions that deforestation occurred on the island well after the arrival of man and that it caused the cessation of the monolithic statue building. There is a general assumption, which can be traced to La Pérouse², that large timbers, obtained from former tall forests, must have been required to transport the statues^{3,4}. We believe that large timbers may not have been necessary for statue building; in particular, we have shown that the Y-sledge and bipod hypothesized by Mulloy⁴ is not very efficient⁵.

The pollen records described by Flenley and King confirm Skottsberg's description of the island's impoverished biota as "waifs and strays"⁶. Only four of the identified genera of "probable trees and shrubs" became extinct on Easter Island before protohistoric times. Flenley and King identified Heliantheae, a tribe of primarily herbs and shrubs, and suggested *Chrysogonum* L. and *Campylothea* Cass. as the probable genera. Some tree-like species do occur in Heliantheae but, because of its herbaceous origin, its wood character is immature and there is little development of secondary xylem^{7,8}. Consequently, the wood is soft and porous. Tree-like species within *Campylothea* are no taller than ~2 m, and within *Oparanthus* Sherff. (based on the former *Chrysogonum* sect. *Quadramera* F. Brown) there is only one known tree, *Oparanthus rapense*, which, although of medium size, is definitely soft⁹. The presence of another former genus, the putrid-smelling *Coprosma*, within the last few thousand years was indicated by only traces of pollen. Of its approximately 100 species of shrubs and trees¹⁰, useful timber can come only from a few of the largest. One can argue on biogeographical grounds that the Easter Island *Coprosma* probably represents a known or extinct species of the *Pyrifolia* group. However, all except possibly one are shrubs or small trees of limited or no practical value. The final missing species identified is palm, probably a *Pritchardia*. Some species of *Pritchardia* can grow to up to 20 m, but 65% are less than 5 m in height (some are even smaller than 1 m)^{11,12}. The useful height of a palm is considerably less than its standing height¹³ and as palm wood is flexible its use as a lever or for shear legs is very limited. Palm logs could have served as rollers, but they are not particularly efficient when used on the ground or on a roughly constructed trackway, especially if they are of small diameter⁵. A further problem is that palm wood is not very durable—the trunks of nearly all species split on drying and,

without preservatives, tend to rot¹³. We believe, therefore, that the trees and shrubs indicated by the pollen analysis cannot be clearly taken as a significant source of large timber suitable for moving the statues.

Flenley and King^{1,14} had a considerable problem in dating their pollen records and expressed reservations about 5 of their 13 ¹⁴C dates. With a single exception, the suspected dates include those most relevant to prehistoric human occupation. Pollen from former genera continues into protohistoric and recent levels at all three sites, indicating a degree of sediment mixing which, at least in part, could have been caused by human activity in the crater and caldera lakes^{15,16}. At Rano Raraku, recent gully erosion, reported elsewhere¹⁴, may also be responsible.

The Rano Kao core provides the most relevant data for the period of human habitation. However, the only two ¹⁴C dates for this core are in conflict. Flenley and King rejected a date of 1,040 ± 60 BP from the middle of the core because of a drop in loss-on-ignition values. It is on the basis of the single ¹⁴C date of 990 ± 70 BP at a depth of ~9 m that Flenley and King argue for deforestation occurring at ~500 BP. We think that it is premature to draw this conclusion without at least one other ¹⁴C date as a check.

If large timbers were important they would have been required throughout the statue-building era, from ~AD 1000 to possibly as late as 1680¹⁷. The largest statue ever erected, the 80-tonne Paro, belongs late in the sequence, yet when Europeans arrived the longest piece of timber seen on the island was only a plank of a little over 2 m (refs 2, 18). Could all large timber artefacts have completely disappeared in such a short interval?

JOHAN KAMMINGA

Division of Prehistory,
La Trobe University,
Bundoora, Victoria 3083,
Australia

BRIAN COTTERELL

Department of Mechanical Engineering,
University of Sydney,
New South Wales 2006,
Australia

1. Flenley, J. R. & King, S. M. *Nature* **307**, 47–50 (1984).
2. La Pérouse, J. F. de G. *Voyage de la Pérouse Autour du Monde* (Paris, 1797; Engl. transl.: Johnson, London, 1978).
3. Heyerdahl, T. *Sea Routes to Polynesia* (Allen & Unwin, London, 1957).
4. Mulloy, W. J. *Archaeol. phys. Anthropol. Oceania* **5**, 1–23 (1970).
5. Cotterell, B. & Kamminga, J. *Cultural Mechanics. An Introduction to the Mechanics of Early Technology* (Penguin, London, in the press).
6. Skottsberg, C. *The Natural History of Juan Fernandez and Easter Island* Vol. 1, 193–438 (Almqvist & Wiksells, Uppsala, 1956).
7. Carlquist, S. *Trop. Woods* **108**, 1–30 (1958).
8. Carlquist, S. *Aliso* **6**, 13–24 (1965).
9. Brown, F. B. H. *Bernice P. Bishop Mus. Bull.* **130** (1935).
10. Oliver, W. R. B. *Bernice P. Bishop Mus. Bull.* **132** (1935).

11. Beccari, O. & Rock, J. F. *Bernice P. Bishop Mus. Mem.* **8** (1921).
12. Caum, E. L., *Bernice P. Bishop Mus. Occ. Pap.* **9** (5) (1930).
13. McQuire, A. J. in *Coconut Wood—1979. The Proceedings*, 24–28 (Philippine Coconut Authority, Manila, 1979).
14. Flenley, J. R. *Asian Perspect.* **22**, 33–40 (1979).
15. Metraux, A. *Bernice P. Bishop Mus. Bull.* **160** (1940).
16. McCoy, P. C. *Easter Island Settlement Patterns in the Late Prehistoric and Protohistoric Periods* (Easter Island Committee, International Fund for Monuments Inc., New York, 1976).
17. McCoy, P. C. *The Prehistory of Polynesia* (ed. Jennings, J. D.) 135–166 (Harvard University Press, Cambridge, Massachusetts, 1979).
18. Cook, J. *A Voyage Towards the South Pole and Round the World* (1772–75) Vol. 1, 2nd edn (Strahan & Cadell, London, 1777).

FLENLEY AND KING REPLY—Kamminga and Cotterell appear to be taking issue with us on two main points: (1) whether the giant statues were moved with the aid of timber, and (2) the date of deforestation. As regards point (1), there is no issue to take. We did not conclude, even tentatively, that the statues were moved in any particular way, but merely that our results did not conflict with those of Mulloy¹, who proposed the use of timber; the results are not conflicting. The pollen results show that *Sophora* and *Palmae* formerly grew on the island more abundantly than in the historic period. We have new evidence, which we hope to publish soon, that the palm was not a species of *Pritchardia*, so the indication by Kamminga and Cotterell that only 35% of *Pritchardia* species exceed 5 m height is irrelevant. As to how the statues were moved, we have no particular opinions.

On the subject of dating, we agree that there are problems. We believe the main reason for the reservations—the possibility of inwashed soil carbon—to be a reasonable one. If, however, the reservations are ignored, one would conclude that deforestation occurred after 6,850 ± 60 BP at Rano Raraku, after 2,100 ± 50 BP at Rano Aroi and both before and after 1,040 ± 60 at Rano Kao. These dates do not conflict with our tentative conclusions (Fig. 4 of ref. 2). As 1,040 ± 60 BP is not statistically separable from 990 ± 70 BP, they do not even conflict with our statement that clearance occurred since 990 ± 70 BP at Rano Kao.

We all agree, I am sure, that more dates are desirable. These have now been obtained and will shortly be published. We agree that it is odd, if clearance was so late, that no very large wooden artefacts survived into historic times. It is worth noting, however, that Palmer in 1868 recorded "boles of large trees, *Edwardsia* (*Sophora*), Coco palm, hibiscus, decaying in some places"¹³. There would have been time for these species to have been introduced after the discovery of the island in 1722, but the pollen record suggests that the *Sophora* and a palm were in fact native. It seems possible that Palmer's "Coco palm" was not coconut but the native

Easter Island palm, whatever that may have been.

J. R. FLENLEY
S. M. KING

Department of Geography,
University of Hull,
Hull HU6 7RX, UK

1. Mulloy, W. J. *Archaeol. phys. Anthropol. Oceania* **5**, 1-23 (1970).
2. Flenley, J. R. & King, S. M. *Nature* **307**, 47-50 (1984).
3. Palmer, J. L. *R. geog. Soc. J.* **40**, 167-181 (1870).

Origin of granulites

THE recent 'statistical' study on granulite-facies rocks by Ben Othman *et al.*¹, in which 32 samples of granulites of many different lithologies and from a wide range of geological environments ranging in age from 2,900 to 200 Myr, have been used to make broad generalizations on the evolution of the lower continental crust and to propose a generalized tectonic model of granulite genesis. Using Nd and Sr model ages, they conclude that, in many cases, internal differentiation of continental crust follows primary differentiation of continental crust from the mantle by a "large time interval". I disagree strongly with this shotgun approach to such a major geological problem, particularly in view of the many published geological, petrological, geochemical and isotopic studies on granulite terranes. I maintain that broad generalizations based on so few data from an indiscriminate grouping of petrogenetically different granulites can be misleading. Quite apart from this, the paper contains several geological and geochronological inaccuracies. Only a few major criticisms can be mentioned here.

The calculation of geologically-meaningful model ages (especially for Sr) requires assumptions about the isotopic evolution of Nd and Sr in the mantle which are too simplistic to apply in a generalized manner to rock units such as granulites which may have igneous and/or sedimentary protoliths of mixed age and provenance, a wide range of crustal residence ages, a complex polymetamorphic history, and highly variable Sm/Nd and Rb/Sr ratios. Careful and precise Sm-Nd, Rb-Sr, Pb/Pb isochron (and U-Pb zircon) age and initial ratio studies in individual terranes, controlled by detailed geological, petrological and structural studies, are preferable to geochronologically-dubious model ages. Although not many individual terranes have yet been studied by all available isotopic methods (particularly by Sm-Nd), the combined approach has already shown that every high-grade metamorphic terrane must be judged on its own merits. In particular, crustal residence time varies widely both within a single terrane and from one granulite terrane to another, suggesting that no generalized tectonic model, such as that proposed by

Ben Othman *et al.*¹, is likely to apply to all granulite terranes.

As an example, Sm-Nd, Rb-Sr, Pb/Pb whole rock, as well as U-Pb zircon, age data on granulites from Finnish Lapland yield concordant ages at ~1,900-2,000 Myr, and show that the time interval between granulite-facies metamorphism and mantle-derived magmatism cannot be resolved². (In this connection, Ben Othman *et al.*¹ refer to a single Lapland granulite whose very discordant model and geological/metamorphic ages bear no resemblance to those given in ref. 2). A similar situation holds in the early Archaean Amitsoq gneisses of West Greenland, where mantle extraction of gneiss protoliths is penecontemporaneous with amphibolite- and granulite-facies metamorphism at ~3,600 Myr, as determined by Sm-Nd, Rb-Sr, Pb/Pb and U-Pb age methods (ref. 3 and N. W. Jones, personal communication). In the Lewisian complex of north-west Scotland, granulite-facies metamorphism and concomitant depletion in incompatible elements may have postdated mantle differentiation by ~200 Myr but almost certainly as part of the same major crustal accretion-differentiation event⁴. (Ben Othman *et al.*¹ wrongly assert that the ~2,700-Myr granulite-facies metamorphism is the age of a major orogenic event associated with granitoid injection. The main period of granitoid injection in the Lewisian complex occurred ~1,000 Myr later⁵.)

In many cases, granulite-facies metamorphism is superimposed on very much older igneous, sedimentary or metamorphic rocks during some major tectonic episode quite unrelated to the primary crustal accretion-differentiation event. Such situations can be resolved by detailed age and isotope studies, using the full array of available methods. Of particular interest here is the Pb/Pb whole rock method, which can date episodes of U-depletion (granulite-facies metamorphism) occurring during, or long after, primary rock formation^{6,7}. Thus in northern Norway, granulite-facies metamorphism at ~1,800 Myr postdates production of juvenile mantle-derived continental crust by ~800-900 Myr (refs 8, 9). Note that several cases have been reported where the Rb-Sr whole rock system is severely disturbed by later tectonothermal events and geochemical open-system behaviour, which renders the calculation of Sr model ages in complex metamorphic terranes even more questionable. To what extent the Sm-Nd whole-rock system responds to such events is not yet well established, but much work is in progress.

The above criticisms are analogous to those levelled at an earlier paper by Allègre and Ben Othman on granitoid rocks¹⁰⁻¹², which employed a similarly simplistic isotopic approach to a vastly complex, variable, polygenetic family of rocks extending over most of geological

history. Just as there are granites and granulites, there most certainly are granulites and granulites.

S. MOORBATH

Department of Geology and Mineralogy,
University of Oxford,
Parks Road,
Oxford, OX1 3PR, UK

1. Ben Othman, D., Polvé, M. & Allègre, C. J. *Nature* **307**, 510-516 (1984).
2. Bernard-Griffiths, J. *et al. Precamb. Res.* **23**, 325-348 (1984).
3. Griffin, W. L., McGregor, V. R., Nutman, A., Taylor, P. N. & Bridgwater, D. *Earth planet. Sci. Lett.* **50**, 59-74 (1980).
4. Hamilton, P. J., Evensen, N. M., O'Nions, R. K. & Tarney, J. *Nature* **277**, 25-28 (1979).
5. Watson, J. in *Geology of Scotland* (ed. Craig, G. Y.), 23-47 (Scottish Academic, Edinburgh, 1983).
6. Moorbath, S., Welke, H. & Gale, N. H. *Earth planet. Sci. Lett.* **6**, 245-256 (1969).
7. Moorbath, S. & Taylor, P. N. in *Precambrian Plate Tectonics* (ed. Kröner, A.) 491-525 (Elsevier, Amsterdam, 1981).
8. Griffin, W. L. *et al. J. geol. Soc. Lond.* **135**, 629-647 (1978).
9. Jacobsen, J. B. & Wasserburg, G. J. *Earth planet. Sci. Lett.* **41**, 245-253 (1978).
10. Allègre, C. J. & Ben Othman, D. *Nature* **286**, 335-342 (1980).
11. Moorbath, S. *Nature* **305**, 73 (1983).
12. Pankhurst, R. J. *Nature* **305**, 73 (1983).

BEN OTHMAN ET AL. REPLY—Moorbath's first criticism involves the study on granulitic facies rocks¹. However, we do not understand the point that he is trying to make. For example, he criticizes our conclusion on age relationships in granulite facies rocks whereas our claim is that it is impossible to derive any such relationship. He further argues that model ages cannot be used instead of direct radiometric ages. This is, of course, true and we make no mystery about that. In fact, the whole purpose of our paper was to compare the one with the other.

Moorbath's second criticism deserves more attention and raises a fundamental problem of scientific methodology. The question is whether it is more appropriate to tackle a given geochemical problem by a statistical study with many rocks of different origins first or by a series of detailed regional case studies? The answer to this has been the key to the strategy of every geochemistry research group. Moorbath's firm answer is that case studies must take precedence. Let us examine this point before stating our own arguments.

The statistical approach (meant to define the study of many samples picked at random from different regions) has had many significant successes in isotope geochemistry. For example: (1) The first study of lead isotopes in galenas by Nier² was performed on 12 galenas from Australia to North America. (2) The age of the Earth was first determined by Patterson³ with four samples. (3) The first study of lead isotopes in feldspars which lead to the first model of continental growth was made by Patterson and Tatumoto⁴ on six feldspar concentrates. (4) The first study of Sr isotopic variations by Hedge and Walthall⁵ was based on data from 10 samples of various origins. (5) The first study of Nd isotopic variations and the first evidence of a correlation between the Nd and Sr isotopic composition of rocks was made using 18 samples

from different regions covering various rock types (from basalts to carbonatites)⁶. (6) the first study of Nd isotopic variations in sediments by McCulloch and Wasserburg⁷ presented data from two different continents with seven samples.

All those pioneering studies involving new concepts and often new techniques not only opened up whole new fields of research but reached conclusions which are still valid. To be more specific, many of the ideas proposed by us (C.J.A. and D.B.)⁸ in our statistical study of Nd isotopes in granitoids have been confirmed by detailed case studies⁹⁻¹⁴. We thus contend that the statistical approach has been a success in isotope geochemistry.

We must stress that we do not favour one approach over the other. Both are not mutually exclusive and it is clear that case studies are the natural and fundamental complements of statistical analysis. It is also clear to us that the statistical study must come first because it gives the framework for further investigations. The danger in this most important matter is that dogmatism may cause the neglect of benefits to be gained from different approaches.

CLAUDE J. ALLÈGRE
DALILA BEN OTHMAN
MIREILLE POLVÉ

Département des Sciences de la Terre,
Universités de Paris 6 et 7,
4 Place Jussieu,
75230 Paris Cedex 05, France

1. Ben Othman, D., Polvé, M. & Allègre, C. J. *Nature* **307**, 510 (1984).
2. Nier, A. O. J. *Am. chem. Soc.* **60**, 1571 (1938).
3. Patterson, C. C. *Geochim. cosmochim. Acta* **10**, 230 (1956).
4. Patterson, C. C. & Tatsumoto, M. *Geochim. Acta* **28**, 1-22 (1964).
5. Hedge, G. F. & Walthall, F. G. *Science* **140**, 1214-1217 (1963).
6. Richard, P., Shimizu, N. & Allègre, C. J. *Earth planet. Sci. Lett.* **31**, 269 (1979).
7. McCulloch, M. T. & Wasserburg, G. J. *Science* **200**, 1003 (1978).
8. Allègre, C. J. & Ben Othman, D. *Nature* **286**, 335 (1980).
9. Hamilton, P. J., O'Nions, R. K. & Pankhurst, R. J. *Nature* **287**, 279 (1980).
10. De Paolo, D. J. *Nature* **291**, 193 (1981).
11. De Paolo, D. J. *J. geophys. Res.* **86**, 10470 (1980).
12. De Paolo, D. J. *Science* **209**, 684 (1980).
13. McCulloch, M. T. & Chappell, B. W. *Earth planet. Sci. Lett.* **58**, 51 (1982).
14. Ben Othman, D., Fourcade, S. & Allègre, C. J. *Earth planet. Sci. Lett.* **69**, 290-300 (1984).

Optical quality during crystalline lens growth

ACCORDING to Fernald and Wright¹, teleost fish eye lenses of radius R have a homogeneous core of refractive index (r.i.) 1.56 and radius $0.67R$ and a gradient index cortex with a surface r.i. of 1.38. Their conclusion is based on the constancy of the focal length to radius ratio (f/r) as layers are peeled off the lens (Fig. 3 of ref. 1), and on the fact that lens cores immersed in Ringer's solution give blurred foci (Fig. 1 of ref. 1). It had previously been proposed^{2,3} that fish lenses had a gradient r.i. throughout the lens, with r.i.

similar to that given by

$$N(r)^2 = N_c^2 + (N_s^2 - N_c^2)(r/R)^2 \quad (1)$$

where N_c and N_s are the indices at the centre and surface of the lens, respectively and r is the distance from the lens centre.

We consider that Fernald and Wright's proposed core model is inconsistent with their observed f/r values and with the well-corrected lens as observed on the cover photo of ref. 1. However, we show that a model having a gradient r.i. throughout the lens is consistent with these data.

To illustrate this we contrast results for: (1) a core model with a cortex r.i. given by equation (1) for $r/R > 0.67$, and N_c and N_s chosen to give core and surface r.i. of 1.56 and 1.38, with (2) a gradient model having no homogeneous core but whose r.i. is given by equation (1). If the lens in (2) is peeled to a radius r and immersed in a medium of index N_w (1.334 for Ringer's), its focal ratio is

$$f/r = 1/2[1 - N(r)N_w/N_c^2] \quad (2)$$

Our gradient model has $N_s = 1.38$ and $N_c = 1.5383$ (compared with 1.53 in ref. 2) and for an intact lens in Ringer's, equations (1) and (2) give $f/r = 2.252$, as observed¹.

The cover photo of ref. 1 shows that the intact teleost lens is well corrected: a pencil limited by an aperture of radius $0.5R - 0.55R$ on the posterior surface has a minimum blur circle of radius $0.02R - 0.03R$ located $2.3R - 2.4R$ from the lens centre. Since the focal length is $2.25R$ (ref. 1) the lens is slightly overcorrected for spherical aberration: the best focus is posterior to the paraxial focus.

We ray-traced the core and gradient models to simulate the cover of ref. 1. The core model is strongly undercorrected for spherical aberration and differs markedly from this photo: the minimum blur circle for a $0.55R$ pencil has a radius $0.2R$ ($10 \times$ that on the cover photograph) and is $1.75R$ from the lens centre (compared with $2.4R$). We varied the core size and refractive index, and the form of the gradient. No model with a homogeneous core, with or without a discontinuous r.i. at the core boundary, gives the well-corrected focus on the cover of ref. 1.

However, the gradient model agrees with the cover photo: the same pencil has a minimum blur circle of radius $0.02R$ at $2.4R$ from the lens centre.

The paraxial f/r was computed for peeled lenses. The core model fails to agree with Fig. 3 of ref. 1 for any value of r (Fig. 1). Fernald and Wright assigned a core r.i. of 1.56 (and refer to Matthiessen, who actually quotes 1.5-1.53²). Whereas a homogeneous lens of index 1.56 has $f/r = 3.45$, to reproduce the observed $f/r = 2.951$ a r.i. of 1.606 is required, considerably higher than values for proteins found in lenses. However, the resulting lens is even more undercorrected.

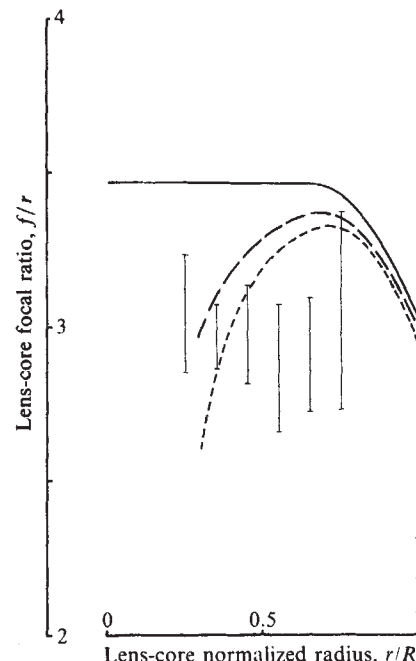


Fig. 1 The focal ratio f/r of a lens with a homogeneous core as a function of the normalized peeled radius r/R for paraxial rays (—), and for pencils of radius $0.2R$ (---) and $0.25R$ (- - -). The error bars are the data from Fig. 3 of ref. 1.

The gradient model agrees with experiment for large r , but not for $r < 0.6R$. A possible explanation is that Fig. 3 of ref. 1 was obtained using a thin laser beam whose radius was not reduced proportionally when the smallest lens cores were examined. When the lens is peeled the index change across the surface increases and the lens becomes increasingly undercorrected for spherical aberration (Fig. 1 of ref. 1). The importance of aberrations when the beam radius is a large fraction of the core radius was demonstrated by ray-tracing. If beam radii are $0.2R$ and $0.25R$, approximating the beams on the cover of ref. 1, the core model still fails to agree with observation (Fig. 1), whereas the gradient model fits quite well (Fig. 2).

Although r.i. discontinuities are a cause of reflections and 'growth rings' observed in crystalline lenses (ref. 1 and Fernald's reply below), these discontinuities may be small and at a cellular or fibre level, so that for optical analysis the r.i. change is equivalent to a continuous gradient. Discontinuities may also be induced as a result of handling¹.

The gradient model for the teleost lens is similar to models for rat lenses. In the rat, evidence for equation (1) came from protein concentration measurements⁴ and from a non-destructive measurement of the r.i. distribution⁵. In the latter the exit angles of laser beams are measured as a function of their incident aperture on the lens and the distribution derived via an Abel integral inversion⁶. This rat model agrees with comprehensive observations of optical performance⁷.

The support of M.C.W.C. by a CSIRO

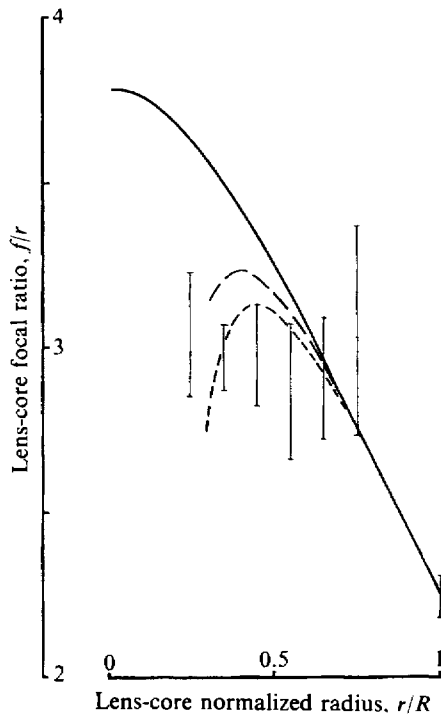


Fig. 2 The focal ratio f/r of the gradient lens model; curves as for Fig. 1.

Australia Postdoctoral Fellowship is acknowledged. We thank Colin Pask for bringing the original paper to our attention and for initial discussions, and W. S. Jagger for access to his unpublished translation of Matthiessen's work.

MELANIE C. W. CAMPBELL*

Department of Applied Mathematics,
R. S. Phys. S. and
Department of Physiology, JCSMR,
Australian National University,
Canberra, ACT, Australia

P. J. SANDS

Division of Computing Research,
CSIRO, Canberra,
ACT, Australia

*To whom correspondence should be addressed. Present address: School of Optometry, University of Waterloo, Waterloo, Ontario N2L 3G1, Canada.

1. Fernald, R. D. & Wright, S. E. *Nature* **301**, 618-620 (1983).
2. Matthiessen, L. *Pflügers Arch. ges. Physiol.* **21**, 287-307 (1880); **27**, 510-523 (1882); **38**, 521-528 (1886).
3. Pumphrey, R. J. in *The Cell and the Organism* (ed. Ramsay, J. A.) 193-208 (Cambridge University Press, 1961).
4. Philipson, B. *Invest. Ophthalmol.* **8**, 258-270 (1969).
5. Campbell, M. C. W. *Vision Res.* **24**, 409-415 (1984); thesis, Australian National Univ. (1982).
6. Chu, P. L. *Elect. Lett.* **13**, 736-738 (1977).
7. Campbell, M. C. W. & Hughes, A. *Vision Res.* **21**, 1129-1148 (1981).

FERNALD AND WRIGHT REPLY — Brewster¹, Maxwell² and Matthiessen³ independently investigated fish eye lenses and each concluded that such a spherically symmetrical lens which produces a reasonably focused image must have a gradient of refractive index from the centre to the edge. More than a century later, Luneburg⁴ formulated an integral equation describing the refractive index profile of a spherical lens which produces a perfect image and solved it for one case in which an exact analytical solution was possible. Since then, such lenses have

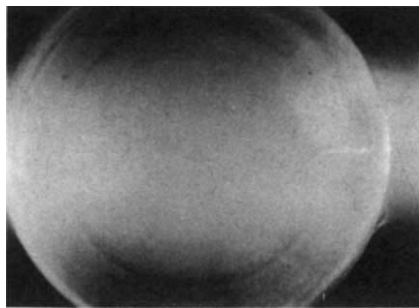


Fig. 3 Lens of the teleost, *H. burtoni*, imaging a laser beam (494 nm) placed about 1 focal length behind it. Note the concentric discontinuities in the lens substance in the cortex of the lens. The innermost discontinuity corresponds to the juncture between the core and cortex of the lens.

become known as 'Luneburg lenses' and the solution extended⁵⁻⁸ so that the refractive index gradient which depends on the focal length of the lens (Fig. 2 of ref. 8) can generally be given in analytical form. Such lenses must produce a perfect image and have a continuous refractive index gradient from the centre to the edge of the lens.

We analysed the teleost fish lens⁹ and hypothesized that the distribution of refractive index diverged from that of the Luneburg lens based on physical measurements of the refractive properties of the intact lens as compared with the lens as it was surgically reduced in size. We suggested that the lens had a core of approximately uniform refractive index, surrounded by a shell with an appropriate refractive index gradient.

Based on ray tracing, Campbell and Sands suggest that our hypothesis is incorrect, offering two possible sources for this difference. First, that the equivalent refractive index we calculate at the core of the lens is higher than Matthiessen³ found, and second, that in measuring the paraxial focal length, our laser beam may not have been reduced in proportion to the size of the lens being measured.

Regarding the first suggestion, measurements more recent than Matthiessen's of the magnitude of the refractive index of the fish lens near the centre, including our own, place the value near 1.56, and for some crystalline proteins, refractive indices of approximately 1.6 have been measured¹⁰. Thus, our computed value of the equivalent refractive index of 1.606, while high, is not unreasonable given the variance in our measurements (Fig. 3 of ref. 9). Regarding the second suggestion, we did reduce the size of our measuring beam in proportion to the core size being measured to the limiting aperture available. So, for larger lenses, we met this condition, whereas for very small lenses we did not. Nonetheless, there are not significant, consistent differences between these measurements as predicted by Campbell and Sands.

We propose that there may be another way to account for the discrepancy between the physical measurements and

the theoretical ray tracing. As noted above, there are two requirements for 'Luneburg' type of refractive index gradient profile to apply: first, the lens must be perfect (without spherical aberration), and second, the refractive index must vary continuously (without discontinuities) throughout its extent. It is not known how deviation from these two conditions would be reflected in deviations from an otherwise obligatory 'Luneburg' refractive index profile.

The fish lens, although not optically perfect, is clearly of high quality (see cover photograph of ref. 9) and our measurements show that the resolution of the lens approaches the diffraction limited value (R.D.F. and S.E.W., in preparation). Thus, the first condition is approximately met. It is not clear, however, that the refractive index is without discontinuity. Figure 3 shows a lens from the African cichlid fish, *Haplochromis burtoni*, which is transilluminated. Distinct concentric discontinuities are evident, particularly in the cortex, as have been reported previously for teleost lenses (see Plate I of ref. 11). Indeed, in many fish, particularly larger ones, a discontinuity in the lens is visible during retinoscopic inspection. The existence of different zones within the fish lens has been demonstrated as optical anisotropy (birefringence)^{11,12} and related to differences in supramolecular organization¹³. This is consistent with the fact that the lens crystallins are differentially synthesized during lens development and therefore not distributed uniformly through the lens^{13,14}. Thus, there may be one or more discontinuities in the refractive index which must be ascertained before ray tracing can adequately be carried out. Since our original technique was not sensitive enough to detect such discontinuities, we intend to measure the refractive index profile more directly either using the backscattered ray method¹⁵ or protein concentration distribution¹⁶, to test directly our hypothesized profile of refractive index within the teleost lens.

RUSSELL D. FERNALD
STEPHEN E. WRIGHT

Institute of Neuroscience,
University of Oregon,
Eugene, Oregon 97403, USA

1. Brewster, D. *Phil. Trans. R. Soc. Lond.* **311**, 311-317 (1816).
2. Maxwell, J. C. *Cambridge Math. J.* **8**, 188-195 (1854).
3. Matthiessen, L. *Pflügers Arch. ges. Physiol.* **21**, 287-307 (1880); **27**, 510-523 (1882); **38**, 521-528 (1886).
4. Luneburg, R. K. *Mathematical Theory of Optics*, 208-213 (Brown University, Providence, 1944).
5. Stettler, R. *Optik* **12**, 529-543 (1955).
6. Morgan, S. P. *J. appl. Phys.* **29**, 1358-1368 (1958).
7. Southwell, W. H. *J. opt. Soc. Am.* **67**, 1010-1014 (1977).
8. Sochacki, J. *J. opt. Soc. Am.* **73**, 789-795 (1983).
9. Fernald, R. D. & Wright, S. E. *Nature* **301**, 618-620 (1983).
10. Doty, P., Geiduschek, E. P. in *The Proteins, Chemistry, Biological Activity and Methods* (eds Neurath, M. & Bailey, K.) 394-460 (Academic, New York, 1953).
11. Pumphrey, R. J. in *The Cell and the Organism* (ed. Ramsay, J. A.) 193-208 (Cambridge University Press, 1961).
12. Burke, P. A., Farnsworth, P. N. & Bettelheim, F. A. *Curr. Eye Res.* **1**, 689-694 (1982).
13. Papaconstantinou, J. *Science* **156**, 338-346 (1967).
14. Piatigorsky, J. *Differentiation* **19**, 134-153 (1981).
15. Chu, P. L. *Elect. Lett.* **13**, 736-738 (1977).
16. Philipson, B. *Invest. Ophthalmol.* **8**, 258-270 (1969).